

Mapping the journey of cancer patients through the health care system. Part 2: Methodological approaches and basic findings

By Jeff A. Sloan, Shannon Scott-Findlay, Anne Nemecek, Paul Blood, Cheryl Trylinski, Heather Whittaker, Samy El Sayed, Jennifer Clinch, and Kong Khoo

Abstract

This is the second in a series of articles from a line of research whose intent was to construct a complete history of interactions with the health care system. This paper provides details of the methods developed to collect and collate the scattered information regarding the event history (trajectory) that a cancer patient experiences in traveling through the Manitoba health care system from one year prior to diagnosis through to two years post-diagnosis. Survival data were obtained through 1994. Basic population data obtained from this work are also presented, including survival information through to four years post-diagnosis. Issues regarding standardized data recording and detail level of clinical events in the chart record are discussed. This part of the research demonstrates that diverse data sources in the health care system can be linked with a high degree of accuracy and completeness of data.

“What happens when a person is diagnosed with cancer?” The first paper in this series (Sloan et al., 2004) described the circumstances under which this question arose and the difficulties encountered when one tries to put together a framework to answer what, on the surface, would seem to be a basic question. This paper summarizes the data accumulation process and development of the basic methodology for pulling together all available information to produce cancer patient histories. Some of the basic findings that were derived from the methodological processes will also be presented.

Manitoba is a province of 1.1 million people with a socio-economic mix of ethnicity, income, age, and education broadly the same as the Canadian population as a whole. The health care system in Manitoba is administered on the same basis as other Canadian provinces. All of the practising oncologists are located in the provincial capital, Winnipeg, which has a population of 650,000 people. The majority of patients travel to Winnipeg for cancer diagnosis and treatment. However, some chemotherapy treatment is provided in rural cancer outreach centres by family physicians, under the supervision of the oncologists.

Data sources

In 1990, the Manitoba Cancer Treatment and Research Foundation (MCTRF) Cancer Registry registered 6,662 new cases of cancer in Manitoba. Approximately 30% (2,000) of these cancer cases were breast, colorectal, or lung disease. These three sites were chosen to reflect a wide diversity of experience in disease trajectory. Breast cancer typically involves good prognosis and lengthy survival. Lung cancer most often presents with a short and acute symptomatic profile. Colorectal cancer tends to be of intermediate disease intensity and duration relative to the breast and lung cases. The term case refers to a singular diagnosis of a primary cancer tumour. A particular patient could have a number of malignancies and, therefore, generate a number of cases (cancer diagnoses). In our final dataset, there are 2,015 cases (malignancies) diagnosed in a population of 1,979 patients.

Study data were gathered from three sources: the MCTRF computerized patient core database (MAXON), data available in the

MCTRF patient records, and the Manitoba Health physician/hospital claims database. The flow chart in Figure One details the process of data abstraction and compilation.

The MCTRF's MAXON database provided detailed information for each case within the study in four segments. The personal identification section contains the demographic data such as age, sex, residence, health numbers, vital status, and date of birth. The Malignancy Abstract (MA) includes the date of cancer diagnosis, histology and morphology of the tumour. The treatment information (TR) section includes the method and date of treatment and health care facility where treatment was performed. Finally, the record location (RL) section designates the record as an MCTRF treatment record (patient being treated at MCTRF) or a report (cancer diagnosis) registered in the Cancer Registry. Each of the four MAXON sources was kept as a separate file initially to maintain data integrity (see Figure One).

Patient clinical charts contained important data not included in any of the computerized databases and, in particular, disease staging information. The complex process necessary to review the charts to produce pathological TNM staging is the subject of the third article in this series (Scott-Findlay et al., 2004). It was also necessary to supplement the available information about treatments given with further detail on follow-up visits and other contacts between the patient and the MCTRF health care system.

Two separate databases were abstracted from the MCTRF patient charts using Statistical Package for Social Sciences (SPSS-PC). One

Jeff A. Sloan, PhD, at the time of the project was Biostatistician at the Faculty of Nursing, University of Manitoba. He is currently Lead Biostatistician at the Mayo Clinic in Rochester, MN. Shannon Scott-Findlay, RN, PhD(c), at the time of the project was Research Assistant/Nurse for Community Cancer Programs Network (CCPN) at the Manitoba Cancer Treatment and Research Foundation (MCTRF). She is currently a Doctoral Candidate, Faculty of Nursing, University of Alberta, Edmonton, AB. Anne Nemecek, RN, is the Previous Program Director of Community Cancer Programs Network at MCTRF, Winnipeg, MB. Paul Blood, MD, at the time of the project was Radiation Oncologist at MCTRF. Currently, he is a Radiation Oncologist at the B.C. Cancer Agency, Victoria, BC. Cheryl Trylinski, HRT, at the time of the project was Data Analyst for the Community Cancer Programs Network. She is currently an Outcomes Associate with the Cross Cancer Institute, Edmonton, AB. Heather Whittaker, HRA, at the time of the project was Director of Records & Registry at MCTRF. Currently, she is Director of Health Records & Privacy Officer, CancerCare Manitoba, Winnipeg, MB. Samy El Sayed, MD, at the time of the project was Radiation Oncologist at MCTRF and is currently a Radiation Oncologist at Ottawa Regional Cancer Centre, Ottawa, ON. Jennifer Clinch, BSc, MA, at the time of the project was Co-Director, Research Analysis at the WHO Collaborating Centre for Quality of Life in Cancer Care. Currently, she is a biostatistician at the Clinical Epidemiology Unit, Ottawa Health Research Institute. Kong Khoo, MD, at the time of the project was a Medical Oncologist at MCTRF and is currently a Radiation Oncologist at the Cancer Centre of the Southern Interior, Kelowna, BC.

file held demographic and disease specifications (referred to as the SPEC file) while the second contained a record for each interaction between the health care system and a cancer patient as recorded by the MCTRF chart (referred to as the TRANS file). We developed a specialized abstraction sheet for this purpose.

The SPEC file includes basic information such as type of cancer, family history, the TNM staging, extent of disease, noncompliance, and treatment delay indicators. A number of extra variables were added in response to specific requests by the oncologists, however, these data will not be presented in this manuscript.

The TRANS file contains an 'event' record for each encounter the patient had experienced with the health care system as recorded by the MCTRF. As well as the type of transaction, the region within Manitoba (health care facility) where the transaction took place and the health care provider who initiated (authored) the event were identified from the chart. For treatment events, additional detail was provided in terms of disease site, as well as treatment intent and modality. The intent of the TRANS file was to produce a detailed patient history for transactions which took place at the Manitoba Cancer Foundation for ultimate merging with the transactions recorded in the MHSC databases.

To complete the patient case history picture, transactions were needed that took place outside the auspices of the Manitoba Cancer Foundation. Manitoba has a unique comprehensive computerized health database available from Manitoba Health. Manitoba Health assigns the family health numbers (MHSC number), and receives reports from all health care agencies within Manitoba. Manitoba Health also assigns unique personal health identification numbers (PHIN). Each patient encounter with the health care system, either through a hospitalization or a physician's claim, is recorded and stored in one of two MHSC databases referred to as the hospital or physician abstracts respectively. The MHSC data were obtained in collaboration with the Manitoba Centre for Health Policy and Evaluation (MCHPE)

by culling the appropriate information from the hospital and physician claims abstracts databases. Roos and colleagues (Roos et al., 1993) have described the organization and accuracy of the Manitoba Health data, have done extensive work at the MCHPE in extracting usable data from the existing sources, and have provided a framework for collating such data.

Inclusion criteria

Determining the actual number of breast, colorectal, and lung cases diagnosed in 1990 was a nontrivial task. The Manitoba Cancer Treatment and Research Foundation's computerized database, MAXON, was the starting point used. A master list of cancer registry numbers (CR#) was generated from the MAXON database. The initial search compiled a list of 2,323 patients. Complexities in the relational nature of the MAXON database arose and were resolved to address issues of double counting and phantom records (unverified cases). Once corrected, we reached a starting point of 2,123 cases in a total of 2,089 patients. The data for the study were collected using tumours rather than people as a basis for analysis because the MCTRF MAXON system is organized and cancer statistics reported by tumour diagnosis, not people. Hence, each patient in our study could account for more than one cancer case.

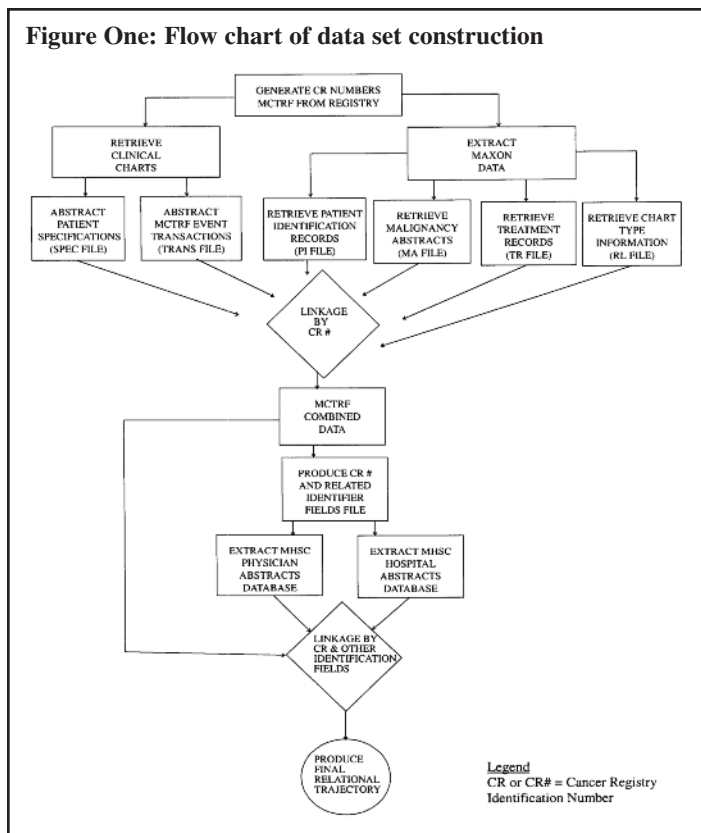
Out of the 2,123 tumours diagnosed in 1990 with the appropriate ICD9 codes, two cases were excluded because the CR numbers were test cases used for systems analysis. Records were also excluded if they were out of province, had an incorrect code (e.g., histology), were military personnel, turned out to have a diagnosis date outside of 1990, or if the MHSC number was missing. The latter condition was imposed to include only cases for which we could combine Manitoba Health data. In total, 124 cases were excluded from the study for these reasons, leaving 1,997 cases.

The 38 cases which did not have MHSC numbers were re-examined due to clinician concerns that we be as complete as possible. Eighteen MHSC numbers were eventually obtained, bringing our total final number of cases diagnosed in 1990 and meeting our eligibility criteria to 2,015. The 2,015 cancer cases diagnosed in 1990 for the three disease sites correspond to 1,979 unique individuals. Thirty-four patients had more than one diagnosis recorded for the year 1990. The 1,979 patients in our study had a total of 2,057 malignancies of all types diagnosed in 1990, representing a multiple malignancy rate of four per cent.

MCTRF data processing and quality control

Processing of the output from the four MAXON datasets and the two datasets transcribed from chart data formed a large part of the team's work. Over one-third (38%) of the 1,979 patients were cared for outside the MCTRF and only had pathology reports from which data could be obtained. MAXON provided information as to whether a detailed clinical chart or only a pathology report was available for a cancer case. Table One indicates that while clinical charts were available for a majority of all cancer site cases, colorectal cases had the largest proportion of pathology report data only. Half of the colorectal cases only had pathology reports compared to 37% for breast and 27% for lung cancers.

Byproducts of the chart review included an investigation into the degree of documentation for treatment noncompliance and delay. An indication of noncompliance with treatment was recorded in only 12 of the 2,015 cases. With the sparseness of data, it is impossible to say whether this is an accurately small number of cases or an underestimate due to poor chart data. One might wish to interpret this small number of noncompliance indications as a measure of noncompliance significant enough to warrant documentation by the clinician. In this sense, the observed number of cases is satisfactorily low. Only 61 cases held charted evidence of a treatment delay. A



broad definition was used to encompass any indication on the chart of a delay in treatment. There is no doubt that at least 61 cases included some form of treatment delay, but we cannot know with any degree of accuracy how many undocumented delays in treatment were actually experienced by patients.

Linkage of datasets

Linkage between the data culled from the MCTRF and the MHSC physician/hospital claims was carried out by the Centre for Health Policy and Evaluation (MCHPE) in consultation with the research team. Of the 1,979 people identified as having been diagnosed, 18 had to be excluded because of invalid or missing MHSC numbers. This reduced the number of people with patient histories to 1,961 individuals.

Records were initially matched simply on MHSC numbers. Subsequently, the extra identification fields were used to produce probabilistic matches for an extra 28 people. In total, this produced 1,926 people whose records could be linked between the MCTRF and MHSC databases. This represents a successful record rate linkage of 98.2%, which compares favourably with other similar matching exercises carried out on the MHSC database.

Once the linkages had been established, all appropriate claims from both the physician and hospital abstract databases were obtained and translated into compatible formats. Extensive manipulation of the physician and hospital abstract datasets was necessary to produce comparable event transaction files using the same format as the TRANS file. These data were then merged with the SPEC, TRANS, and MAXON data to produce complete patient histories for each one of the 1,979 people diagnosed in 1990.

Findings

Overall, an average of six event transactions (median of three events) were culled for each cancer patient from the available MCTRF chart data. The completeness of the chart data had an impact on the number of transactions that were obtainable for

Cancer site	Chart available	Path. report only	Total
Breast	403 (63%)	240 (37%)	643
Colorectal	330 (50.4%)	325 (49.6%)	655
Lung	499 (73%)	182 (27%)	681
Total	1,232 (62%)	747 (38%)	1,979

Cancer Site	<50 years	50+ years	Total
Breast	122 (19%)	521 (81%)	643 (32%)
Colorectal	41 (6%)	614 (94%)	655 (34%)
Lung	37 (5%)	644 (95%)	681 (34%)
Total	200 (10%)	1,779 (90%)	1,979

Cancer site	Two- year survival	Five-year survival
Breast	94%	91%
Colorectal	62%	54%
Lung	29%	21%

each patient. In total, there were 205,107 physician claims for the people who were diagnosed in 1990 with one of the three cancers under study, representing an average of 112 claims per person for the period 1989-1992, or an average of 28 claims per year. The physician claims included 65,169 (32%) for ambulatory visits. Out of province claims totalled 561 (1%). Of all claim codes, 7% were for examinations by specialists and 21% for general examinations.

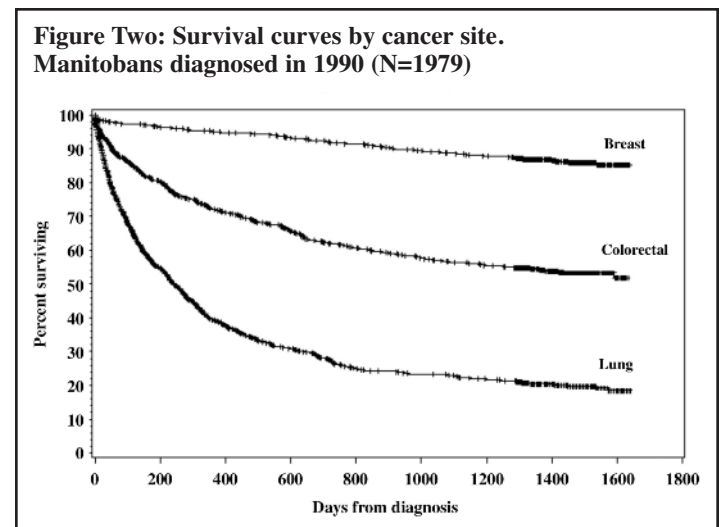
In total, 1,897 hospital admissions were recorded for 1,901 patients out of the 1,979 people diagnosed with breast, colorectal, and lung cancer in Manitoba for 1990 from one-year pre-diagnosis through two years post-diagnosis. Over this period, patients averaged a total of 20 days in the hospital for care directly related to cancer care. Breast cancer patients spent less than half the total number of days than patients with colorectal or lung cancer (10 days versus 24 and 28 days respectively). Median lengths of stays for the three sites were eight, 15, and 17 days respectively. When non-cancer-related hospital admissions were added to the above results, the relative amounts among the disease sites stayed consistent, but the average stays increased to 27 days, 50, and 48 days respectively for breast, colorectal, and lung cancer patients. The greatest amount of time a cancer patient spent in the hospital in the four-year period of observation was 856 days (2.3 years).

Patient demographics

The majority of cancer patients for colorectal and lung sites were male (55% and 62% respectively). Virtually all of the breast cancer patients were female (two males were included). Females presented in greater numbers than males with late stage (stage III or IV) colorectal cancer. In total, 90% of all patients diagnosed in 1990 were over 50 years of age. At the time of diagnosis, this cancer patient population was in their late sixties on average (mean of 68, median of 69 years, standard deviation of 12.8 years). Table Two presents more information on the age of the cancer patient population. These data are comparable to age of onset data published by Statistics Canada for 1990 indicating an average of 70 years of age for the onset of breast and lung cancer (National Cancer Institute of Canada, 1996).

Survival data

The survival curves in Figure Two show the marked differences among the three disease sites. Roughly half of the cancer patients diagnosed in 1990 for the three sites died before June 30, 1994. The more advanced the disease, the less likely an individual was to have survived to the censoring date. The vast majority of breast cancer



patients (86%) were still alive at this date. Just over half of the colorectal cancer patients survived to this point, while only one in five of the lung cancer patients were still living. In particular, differences among the two- and five-year estimated survival rates are given in Table Three.

The hazard functions in Figures Three and Four indicate a monotone low risk of death from breast cancer while colorectal cancer patients experienced a moderately greater risk in the first year post-diagnosis before experiencing a risk comparable to that of the breast cancer patients. Lung cancer patients experienced a dramatically higher risk of death for the first three years post-diagnosis.

Discussion

This was the first study in Manitoba that identified a cohort of cancer patients and linked their registry and cancer chart information with the Manitoba Health data. A major accomplishment of this study was to demonstrate that data from disparate sources and styles of cancer care data could be successfully combined into a comprehensive history of each patient's interaction with the Manitoba health care system. Linkage of 98.2% of the initial population frame instills a high degree of confidence in the quality of the resultant data.

The richness of clinical care can never be totally captured in a statistical dataset. Variables collected in this study do not tell some of the more subtle nuances of care, nor do they include all the variables that the clinicians on the team might have wanted to include. Our data only notes contacts with physicians and related hospitalizations. Cancer patients have contacts with nurses, clinical counsellors, lab technicians, and radiotherapists among others, all of whom contribute to the cancer experience. A balance had to be struck, however, between what could be gathered and what was essential to be gathered. One may disagree with our particular taxonomy of event

classifications, but the taxonomy used was developed with this balancing of competing interests in mind.

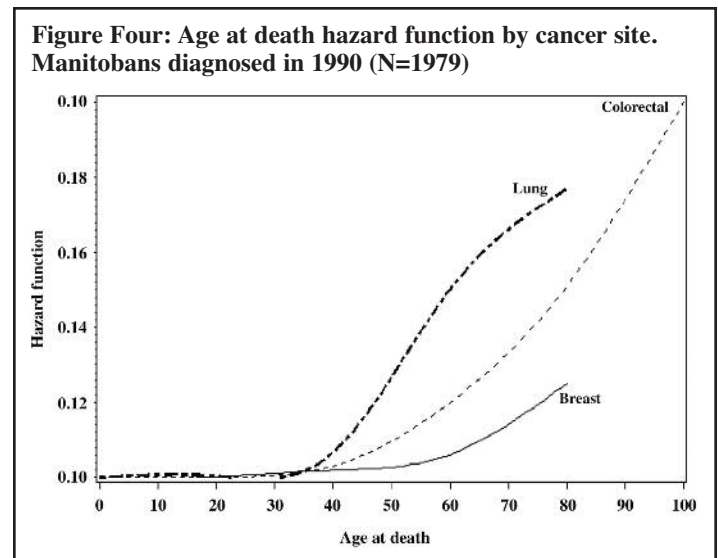
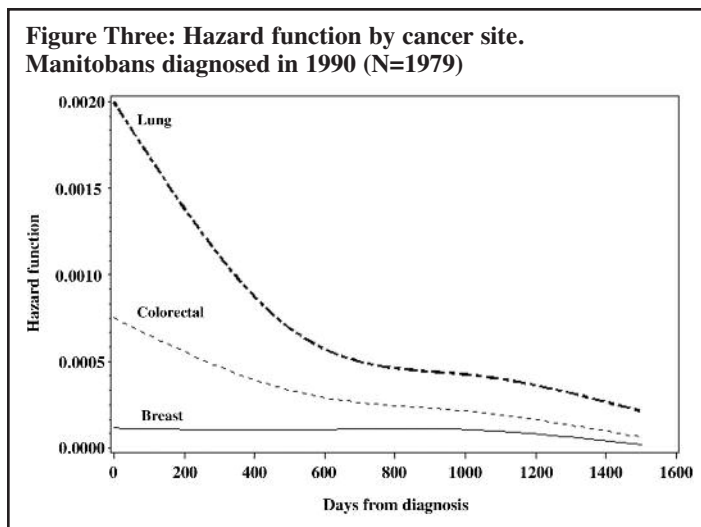
The degree of difficulty and amount of effort required to obtain the necessary charts/reports was formidable. Waiting times for charts were sometimes months long. For example, the final merge of the dataset was delayed a month while waiting for seven charts that had seemingly become lost in the system. It was only after extraordinary search procedures were undertaken that the final charts were located. Subsequent researchers should be cognizant of this fact so that sufficient time is built into the chart abstraction process to allow for such delays. The fact that all charts were eventually located is a credit to the Cancer Registry and the research team's tenacity. Similar studies in London, UK have documented a missing chart rate up to 47% (Vickers & Pollock, 1993).

Through careful consideration of data quality, availability, and operationalization, the study team managed to produce patient histories with complete linkage of all data for more than 95% of all people diagnosed with breast, colorectal, or lung cancer in Manitoba in 1990. This represents the most complete picture to date of the various trajectory paths taken by individuals through the Manitoba health care system from one year prior through four years after a diagnosis of cancer.

Acknowledgements

Financial support for this study was provided by the Manitoba Cancer Treatment Research Foundation and a grant from the Manitoba Medical Service Foundation.

At the time of this study, the Manitoba Oncology Centre was called the Manitoba Cancer Treatment and Research Foundation. It is at present called CancerCare Manitoba.



References

- National Cancer Institute of Canada [NCIC]. (1996). **Canadian cancer statistics 1996**. Toronto: Statistics Canada.
- Roos, L., Mustard, C., Nicol, J., McLerran, D., Malenka, D., Young, T., & Cohen, M. (1993). Registries and administrative data: Organization and accuracy. **Medical Care**, 31(3), 201-12.
- Scott-Findlay, S., Sloan, J., Nemecek, A., Blood, P., Trylinski, C., Whittaker, H., El Sayed, S., Clinch, J., & Khoo, K. (in press). Mapping the journey of cancer patients through the health care system. Part III: An approach to staging. **Canadian Oncology Nursing Journal**.
- Sloan, J., Scott-Findlay, S., Nemecek, A., Blood, P., Trylinski, C., Whittaker, H., El Sayed, S., Clinch, J., & Khoo, K. (2004). Mapping the journey of cancer patients through the health care system. Part I: Developing the research question. **Canadian Oncology Nursing Journal**, 14(3), 183-186.
- Vickers, N., & Pollock, A. (1993). Incompleteness and retrieval of case notes in a case note audit of colorectal cancer. **Quality in Health Care**, 2, 170-174.