Computerised data collection for a cancer centre: A Queensland perspective

IAN FRANKLIN, QUENTEN WALKER AND LEE TRIPCONY

Ian Franklin is a Senior Radiation Oncologist in the Queensland Radium Institute. Quenten Walker is Associate Professor in the University of Queensland and a Senior Radiation Oncologist in the Queensland Radium Institute. Lee Tripcony is the Statistical Officer in the Queensland Radium Institute.

Abstract

The Queensland Radium Institute has had a computerised database for 15 years but there was little uniformity in the data collected and there was considerable scope for improvement. In 1991 the institute decided to upgrade and standardise its data collection system. The improvements will allow easier access to the data for research and audit purposes. The process was effective and the implementation straightforward because of effective planning and training. This paper outlines the process used in the review and implementation of the new database system and describes the data sets in some detail.

Introduction

With the increased demand by consumers for outcome data in the health industry, there is an increasing need for uniform computerised data collection to allow easier comparison of results from different centres. In the United Kingdom the College of Radiologists is developing the Clinical Oncology Information Network (COIN – Karp & Squire 1994), and Mt Vernon Hospital has reported a system which suits their needs (Elespe et al. 1994). In common with most radiation oncology departments, the Queensland Radium Institute has kept statistics on patients referred and treated. These statistics were published annually since the institute was founded in 1944 until it was disbanded with the coming of regionalisation in 1991. In addition, detailed outcome and survival

data have been published annually. This systematic reporting of outcome data, a legacy of Manchester, is, to our knowledge, unique in health care facilities in Australia.

For many years the Queensland Radium Institute patient data was collated manually and consisted mainly of administrative data but included diagnosis, stage and treatment aim as well as the status on discharge at 10 years or death for all patients seen. With computerisation in 1981, the database was extended to include the ICD-9-CM site code and histology. In addition, a much more comprehensive set of clinical data on several diseases has been developed by consensus of the medical staff over the past 15 years. Unfortunately, there was little uniformity in the information collected between different disease types, consistent methods were not always applied to the data collection and updating was often carried out in an ad hoc manner. In some instances it was not possible to relate the data sheet to the patient record for further updating or checking of information. Although these statistics are better than those produced by any other radiation therapy department in Australia (Denham, Hamilton & Joseph 1991), there is still scope for improving the range of data analysed. The Queensland Radium Institute system has many similarities with that of the Mt Vernon Hospital, but has some advantages which suit our practice well and, we believe, make the system easily transferable between centres. This paper outlines the process used in the review and implementation of the new database system and describes the data sets in some detail.

Methods

In 1990 it was decided to set up a committee to assess the ways in which our system could be improved and expanded in order to make the outcome statistics more relevant. It was decided that a uniform and universal policy should be adopted and over the next 18 months the policy was formulated. The administrative data set was unchanged, but it was decided that a standard disease data set, treatment data set and outcome data set would be collected on *all* patients registered at the Queensland Radium Institute. A data set for morbidity as well as one for second courses of treatment are yet to be implemented. Each of the previously designed specific disease data sets was examined and incorporated into the new system with no loss of data, although a few required a small subsidiary data set to be appended. Figure 1 shows the administrative data set and Figure 2 shows the disease, treatment, outcome and relapse data sets.

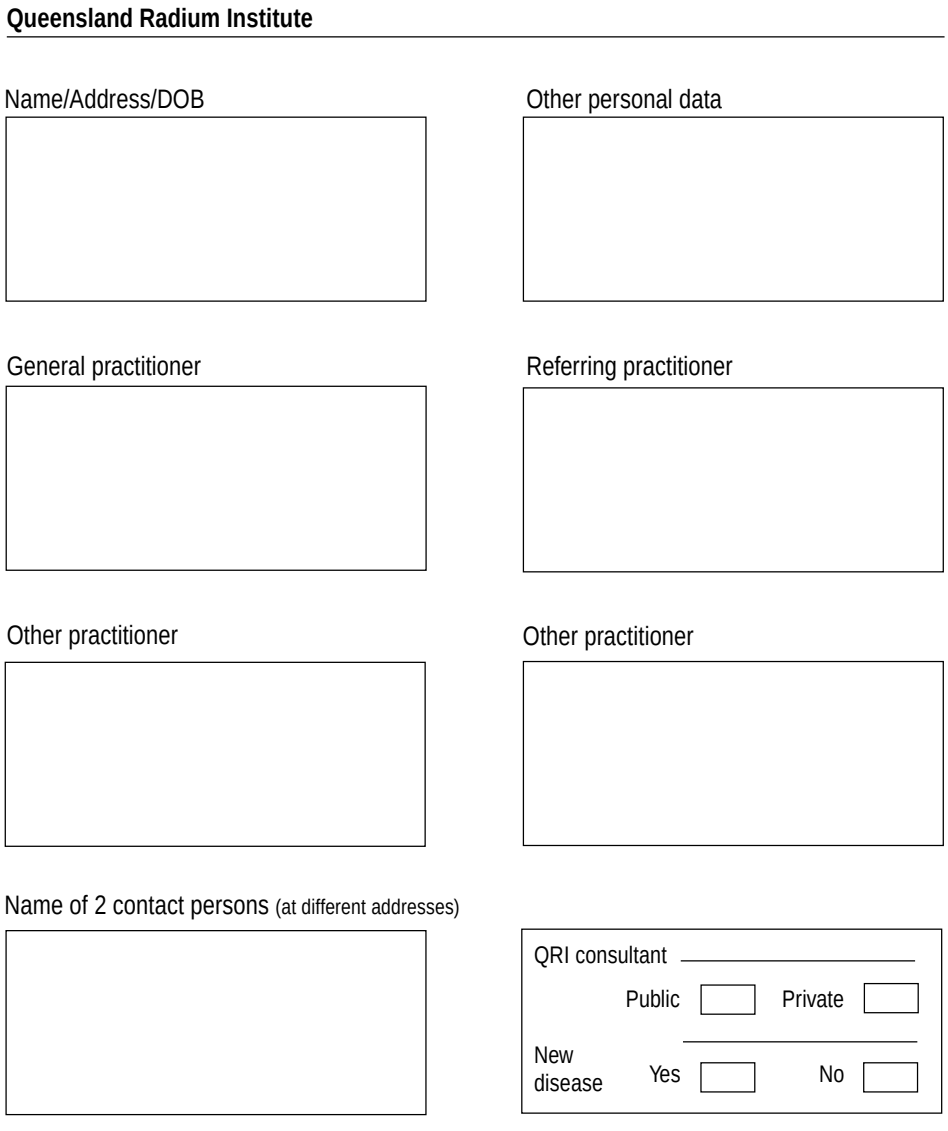

Figure 1: A sample administrative data set

Disease profile all patients

1. Diagnosis (write) _____

2. Site classification
ICD9.CM

3. Date of diagnosis ____ / ____ / ____

4. Basis of diagnosis
1 = Clinical
2 = Markers
3 = Cytology
4 = Histopathology
(Enter highest appropriate number)

5. Histology (write) _____

6. Classification **M**
IDC9.CM
(Refer to Code Sheet)

7. Differentiation (circle number)
0 (N/A) 1 (Well) 2 (Mod) 3 (Poor) 4 (ANAP) 9 (U/K)

8. Staging New disease
(insert number + letter)
T M N

Overall stage
X = unknown or N/A

9. Extent of disease Not new disease only
1 = No
2 = Yes
9 = Unknown
Primary
Regional Nodes
Metastases

10. Treatment profile
Aim
1 = Cure
2 = Palliation
3 = F/U of previously Rx patients

11. Performance status
0 = Fully active, asymptomatic
1 = Ambulatory, capable of light work
2 = Immobile < 50% of time
3 = Immobile > 50 % of time
4 = Bed ridden

12. Date of start of definitive treatment of *this* event ____ / ____ / ____

13. Treatment initiated at QRI
1 = No
2 = Yes

Figure 2: Disease, treatment, outcome and relapse data sets

Figure 2 (continued): Disease, treatment, outcome and relapse data sets

Disease data set

Patients are divided into two groups: those with 'new disease' and those with 'not new' disease. 'New disease' patients include all patients who present to the Queensland Radium Institute with a newly diagnosed neoplastic disease, benign or malignant, or a non-neoplastic disease for an opinion or treatment. They may have had part of their initial treatment elsewhere. This group also includes those patients who have had treatment at the Queensland Radium Institute for a different disease in the past and now present with a second new disease. 'Not new' disease patients are those whose initial treatment management decision did not involve the Queensland Radium Institute in any way and are usually referred with recurrence or metastasis for treatment. Some are patients who have moved to Queensland, having been treated elsewhere, and require continued follow-up.

In addition to the ICD-9-CM site code, we have added the ICD-9-CM histology codes and the basis of diagnosis (clinical, markers, cytology or histology). For new disease patients, the TNM or AJC staging system is used where appropriate. If neither is appropriate, then other well-known staging systems are used, such as FIGO for gynaecological disease and Ann Arbor for lymphoma. For not new disease patients in whom staging is obviously not possible, the extent of disease is recorded as present or not at the primary site, regional lymph nodes and metastatic sites.

This section of the data set is completed by the specialist treating the patient. To add to the accuracy and consistency of the coding, a code sheet appropriate for that disease is included in the patient record at the time of presentation. This code sheet covers the ICD-9-CM codes for that site, the likely histology for diseases in that area and the appropriate staging system. Books containing all of the code sheets are available in all of the consultation suites and planning rooms as well as all doctors having their own copies. It is the doctor's responsibility to ensure that the disease data set is completed in a timely and accurate manner, but they are aided in this task by the clerical staff using reminders when case sheets are overdue.

Treatment data set

Treatment intentions fall into two main categories: cure and palliation. It is, however, necessary to have a third category for those who are referred for ongoing follow-up of disease previously treated elsewhere. Cure is defined as the treatment intention to eradicate all known disease, even if the chance of doing so is small. This applies to the intention on the first day of treatment. If, during a curative course of treatment, the patient develops signs of metastatic disease, then the

original intention was cure and should remain unchanged; however, it will be necessary to complete a relapse data set as well. Palliation refers to treatment which has no chance of permanently eradicating all known disease.

Recorded in this area of the data set is the date of starting this treatment, the ECOG status at the time, the type of treatment and the order in which treatment was given for combined modality treatments. It is also noted if the treatment decisions were initiated at the Queensland Radium Institute.

Outcome data set

The Queensland Radium Institute provides cancer treatment services to the whole of Queensland as well as much of northern New South Wales and some of the Northern Territory. It is not possible to obtain detailed response and outcome data on all patients. In order to obtain appropriate outcome data, various follow-up strategies are adopted as outlined in the algorithm in Figure 3. Detailed data are collected on all patients seen at the Herston and Mater centres of the Queensland Radium Institute for follow-up; details of relapse are obtained on all patients treated with curative intent. All patients will have data regarding date of death or discharge and their disease status at that time. The clerical staff are very rigorous in obtaining this information from the local doctor, the registrar-general's returns or the patient's relatives.

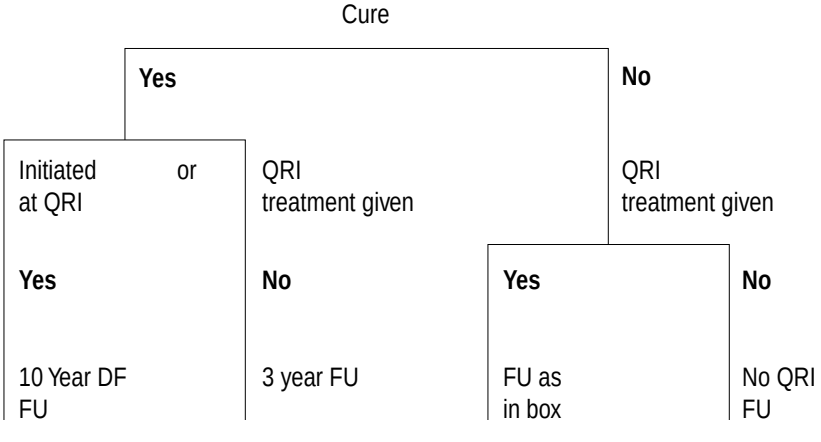

Figure 3: Follow-up algorithm

The patient's record is automatically forwarded to the treating doctor at the end of a course of treatment for coding of the disease and treatment data sets. The outcome data set is automatically flagged for completion six months after the completion of treatment and the patient record will be submitted to the treating doctor again for completion. The flag remains in position until appropriate action has been taken and the data are entered.

These sections are completed by the treating Queensland Radium Institute specialist. If there are gaps in the data entered, the record is returned to the treating doctor for completion. Quality assurance of the data is validated by way of random checks of the database.

Treatment summary

At the completion of each course of radiation therapy, the treatment is summarised by a senior radiation therapist and then validated by the treating radiation oncologist before data entry by the clerical staff. Each course of treatment is summarised in a like manner. There are similar summary forms for use in brachytherapy treatment, gynaecological and non-gynaecological, as well as unsealed sources such as I^{131} and P^{32}. The treatments using isotopes are summarised by the physicists. Figure 4 shows a treatment summary sheet.

A similar summary system is being investigated for treatment with cytotoxic chemotherapy.

Discussion

The development of this system was a significant administrative challenge since most of the members of the institute staff are involved in some aspect of the data collection or data entry (radiation oncologist, physicist, radiation therapist, statistician and clerical staff). All groups had to be educated in the need for the new system and the system had to be modified and refined until it was acceptable to all groups. Once finalised, the staff groups had to be trained in using the system and made aware of its importance and value. It is a great tribute to the commitment of the staff that they have all been prepared to do their own part in the coding, summarising and entry of data. Once the system had been designed, it was necessary to modify the computer programs to accept the new format and to be certain it was functional before transferring to the new system. Those involved in form filling and data entry all had to be trained in the new system before it could be launched. Those goals were achieved with remarkably

Summary of treatment Radiographer sig: H.Scott Date: 10 / 4 / 95 Dr Initials: I.Franklin

Volume treated	Date started	Date finished	Cumulative dose (G_y)	Point of prescription	Maximum dose	No. of fraction	Days	Av. fields per	No. fields in plan	Mode
Large SPIN	10.4.95	10.4.95	10	DEP	10	10	1	1	1	x
Large boost										
Large boost										
Large boost										

Figure 4: A treatment summary sheet

little difficulty by ensuring adequate discussions with appropriate supervisors and relevant in-service training for the staff.

There were some teething difficulties initially, but these have been resolved. There still remain some difficult situations where some aspect of the disease data set is unclear. One of the authors (IF) has been designated the arbiter of problems in the expectation that there will be consistency in the decisions made.

Obviously, to be of major use to the Queensland Radium Institute, the system needs maturity – it has only been operational since 1991. All of the data are collected prospectively and the single-sheet format has made data entry more user-friendly and compatible with a busy clinic. Thus specific prospective studies will be able to be done more easily since the major system is in place. For example, at present we are collecting data in a subsidiary data set on the role of mammography and a false negative result in the delay in initiating treatment for breast cancer to follow up previous retrospective publications in this area (Walker & Langlands 1986; Walker, Gebski & Langlands 1989).

The main advantage of the new system is that it allows much more ready access to data for research purposes. The database can be queried using an appropriate 4QL to enable specific combinations of disease and treatment conditions to be selected accurately. This ensures that all patients satisfying those conditions, and only those patients satisfying those conditions, are selected in that group. The following are examples of various groups not previously available.

1. Rarer conditions which could not be readily accessed before, that is, adenocarcinoma of cervix or SCC of breast. In the past it would have been necessary to look at all cases of the disease and select those of interest. This was a daunting task, with the result that it was rarely done either for internal research or publication.

2. Specific stages of common conditions such as N2 lung tumours or T3 larynx tumours are more easily extracted by having site, histology and stage available in the database.

3. Diseases treated by specific methods such as curative radiotherapy alone in the treatment of laryngeal tumours.

4. Review of patterns of practice with various treatment modalities such as palliative radiotherapy in prostate cancer.

Whilst not obviating the need to study the patient's records, the new system will make it easier to extract only the relevant records and to ensure that all of these relevant records have been extracted.

The system continues to run smoothly because the medical staff see the value of collecting the data, it is departmental policy that the data be collected and this policy is strongly supported by the Director who is involved in completing the data sheets on his patients. There is a mechanism that the patient record will be returned to the medical staff for all the relevant data to be entered on the sheets if this has not been done in a timely manner. This ensures that the database is accurate and up to date.

The system should aid casemix funding and diagnosis related group (DRG) classifications when they are introduced into the division. The medical specialists are well versed in the classification of disease which is part of their daily practice and many of our procedures are also coded at present. The introduction of DRG coding should be accomplished with relative ease and with none of the anguish likely to be experienced by those divisions which are not currently coding their activities.

The details of the hardware are as follows. Initially the system was run on a PDP/11 before it was upgraded to a DEC micro Vax 3400 in 1989. A further upgrade to a pentium has recently taken place, with a dramatic improvement in response time.

Summary

A computerised data collection system has been introduced within the Queensland Radium Institute which will enable greater use to be made of the vast amounts of information collected on our patients. The analysis will assist in improving the management of patients as well as planning for our future needs. The development of the system was achieved by good planning, followed by adequate consultation with the stakeholders. The introduction was smooth and trouble-free because of effective training given to all affected groups and their being made aware of the importance of the project. The system continues to run effectively because the staff own the project and are aware of the advantages to themselves and the Queensland Radium Institute. The full value of the system and the improvements will be realised as more staff analyse the data for research purposes and to audit our clinical care. These analyses will allow the department to publish papers which will advance the contemporary practice of oncology in Australia.

Acknowledgements

We wish to thank the Director of the Queensland Radium Institute for his permission to publish this report. We also thank all of our colleagues, without whose cooperation the database would be useless.

References

Denham JW, Hamilton CS & Joseph DJ 1991, 'Patterns of care studies in Australia', *Australasian Radiology*, vol 35, p 205.

Elespe AM, Dische S, Saunders MI & Pepperell M 1994, 'A computerized data collection system for a cancer centre', *Clin Oncol*, vol 6, pp 237–41.

Karp SJ & Squire CJ 1994, 'Clinical Oncology Information Network (COIN)', *Clin Oncol*, vol 6, pp 5–6.

Walker QJ, Gebski V & Langlands AO 1989, 'The misuse of mammography in the management of breast cancer revisited', *Med J Aust*, vol 151, pp 509–12

Walker QJ & Langlands AO 1986, 'The misuse of mammography in the management of breast cancer', *Med J Aust*, vol 145, pp 185–7.