

Colorectal cancer registration: the central importance of pathology

M D Manser, D F Levine, D F H Pheby, R W Pitcher

Abstract

Background—Changes in cancer care have increased the importance of cancer registries in monitoring trends and outcomes. Registries are increasingly using computerised systems, such as patient administration and histopathology, as data sources. Omissions by registries can cause interpretation errors, but use of multiple data sources can overcome this.

Methods—Registrations of new colorectal cancers in Cornwall were compared with cases identified from primary sources over one year.

Results—Two hundred and thirty cases were identified locally, 93% in documentary records, 89.6% via histopathology, and 81.3% in the clinical data capture module of the patient administration system. Two hundred and forty four cases were known to the regional registry, but after eliminating wrongly assigned and unconfirmed cases only 201 remained. Twenty nine cases identified locally, particularly cases of advanced disease, were unknown to the registry.

Conclusions—District registers based on histopathology augmented from other sources would provide more accurate and less biased information than existing regionally based methods.

(*J Clin Pathol* 2000;53:875-877)

Keywords: colorectal cancer; cancer registration; district registers

Changes in the organisation and delivery of cancer services have increased the importance of monitoring cancer in the population and evaluating the outcomes of cancer care. The Calman-Hine report¹ underlined the role of cancer registries in providing such information. Cancer registries traditionally depended on data from clinical case notes and other documents, but increasingly they make use of computerised systems, such as patient administration and laboratory systems. The decennial review of cancer registration (Alberman report)² recommended the development of links with histopathology laboratories, and Codling and colleagues³ demonstrated a practical way of achieving this within the hospital for cancer registration purposes.

Inevitably, cases are omitted from cancer registry data, whatever sources are used. Measures of ascertainment frequently used in registries include comparison with independent registers, mortality and historical incidence data, as well as capture-recapture methods and registration/mortality ratios.⁴ Various studies have reported ascertainment values varying from around 50% up to 95.8%,⁵ but there have

been few attempts to determine the ascertainment value by direct comparison with primary documents, although Nwene and Smith made such a comparison in 1982.⁶

Methods

To investigate the value of various data sources for cancer registration, we compared registrations of new colorectal cancers in Cornwall with cases identified from primary sources over one year. A research assistant with extensive experience of clinical coding, who also worked in clinical audit as an audit analyst, carried out this work.

All histopathology reports of colorectal cancer/adenoma coded by pathologists and dated 1993 were extracted from the histopathology database using the SNOMED (2nd edition) topography codes T67... or T68... and morphology code M814./. These were then checked against the clinical notes to confirm a definite diagnosis of colorectal cancer.

Manual oncology records for 1993 were also systematically searched for a written diagnosis of colorectal cancer. Lists of patients with a diagnosis of colorectal cancer in 1993 were requested from the radiology department and the stoma nurses. These were then checked against the clinical notes to confirm cases newly diagnosed in 1993. All death registrations were checked using the local death register, to identify notifications to local registrars of births, deaths, and marriages. Deaths from colorectal cancer were extracted for 1993, 1994, and 1995, and checked against the clinical records to identify those first diagnosed during 1993. Records in the clinical data capture (CDC) module of the patient administration system were selected using the ICD-9 codes 153, 154, 211.4, 235.2, and 239.0, and reviewed to confirm their accuracy. Cases where the clinical notes were unclear were discussed with a consultant physician with a special interest in colorectal cancer.

Results

The results of the study are illustrated in fig 1, which demonstrates the extent of shortfalls in

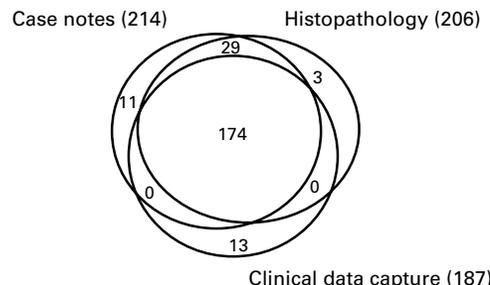


Figure 1 Colorectal cancer cases in Cornwall, 1993: sources of information.

Clinical Audit
Department, Royal
Cornwall Hospitals
Trust, Treliske
Hospital, Truro
TR1 3YL, Cornwall,
UK
M D Manser

Royal Cornwall
Hospitals Trust, West
Cornwall Hospital, St
Clare Street, Penzance
TR18 2PF, Cornwall,
UK
D F Levine

Unit of Applied
Epidemiology,
University of the West
of England, Frenchay
Campus, Coldharbour
Lane, Bristol
BS16 1QY, UK
D F H Pheby

Royal Cornwall
Hospitals Trust,
Treliske Hospital,
Truro TR1 3LJ,
Cornwall, UK
R W Pitcher

Correspondence to:
Dr Levine
dftreans@hotmail.com

Accepted for publication
19 April 2000

Table 1 Ascertainment of colorectal cancer in Cornwall: comparison of data sources

Primary data source	No. cases identified	% Total no. cases	No. cases corroborated from other data sources			No. cases not corroborated
			Documentary sources	Histopathology	CDC	
<i>Documentary sources</i>						
No. cases	214	93.0		203	174	11
κ Score (primary data source <i>v</i> other sources)				0.878	0.540	
No. pairs examined				217	227	
<i>Histopathology</i>						
No. cases	206	89.6	203		174	3
κ Score (primary data source <i>v</i> other sources)			0.878		0.609	
No. pairs examined			217		219	
<i>CDC</i>						
No. cases	187	81.3	174	203		13
κ Score (primary data source <i>v</i> other sources)			0.540	0.609		
No. pairs examined			227	219		

CDC, clinical data capture.

ascertainment in the various data sources examined, and of the overlaps between them. In table 1, these relations are examined in detail. It indicates the extent to which ascertained cases in each data source were confirmed from other sources, the proportion of validated cases identified from each source, and κ scores for consistency derived from pairwise comparisons of the various data sources. The total number of validated cases identified from local sources in Cornwall was 230. Of this total, 93% were found in documentary records, 89.6% in the histopathology computer system, and 81.3% in the clinical data capture module of the patient administration system.

Documentary sources and histopathology were much more effective in identifying cases than CDC, the main data source for the regional cancer registry. Of 258 cases identified in CDC, 71 (28%) were found on clinical review to have been wrongly assigned by coders, in comparison with the documentary evidence in the clinical case notes, usually because of errors in diagnostic coding or in year of diagnosis, leaving just 187 cases. The regional cancer registry identified 331 cases for the whole county. Eighty seven of these proved to be from outside the study area, so the real figure for comparison was 244. Superficially, this appears not dissimilar to the number of cases identified locally, but its composition was very different. Forty of the regional registry cases were not confirmed locally, and had identical errors to those in CDC (including 22 wrongly assigned to colorectal cancer). Three more were unconfirmed. Twenty nine cases identified from local sources were not known to the regional registry. Even after eliminating "out of area" registry cases, the κ score for consistency between the regional registry and the locally compiled list was only 0.472 (total pairs examined = 273). It appears that the registry underestimated the proportion of cases with more advanced disease. Although the overall number of cases known to the registry was comparable with that determined locally, the difference in composition rendered analyses involving subsets of the data (for example, stage based survival) problematic.

Discussion

A Wessex study⁷ suggested that histopathology reports were of insufficient quality to justify

routine use for cancer registration, and indeed could undermine data quality. Our study, on the contrary, suggests that cancer registry data quality could be greatly improved if based, at district level, on histopathology data, augmented by linkage of records from other sources. This has the additional advantage that histopathology reports also allow access to staging data for the registry, and this benefit is likely to be enhanced by the adoption of the recently published minimum data set for the reporting of colorectal cancer surgical specimens.⁸ If our findings are of general application, computerised hospital patient administration and related information systems would appear to be a much less reliable source of data for cancer registration purposes than either clinical case notes or histopathology records; in agreement with these findings, Comber⁹ has pointed out that studies based solely on hospital clinical records may underestimate the true population prevalence of colorectal cancer by around 15%. Where data from computerised hospital information systems are used, they should be augmented by data from more reliable sources.

An additional benefit of the use of data from multiple sources is the creation of opportunities for data validation, leading to improvements in data quality, which should in any case be routinely monitored using standard measures.¹⁰ These include capture-recapture methods, which have been used in recent studies of comprehensiveness of ascertainment by cancer registries as far apart as New Zealand and Cuba, although there may be problems with such methods where there are appreciable numbers of false positive and/or false negative diagnoses in one of the two or more data sources being compared.

Other recent studies of the comprehensiveness of case ascertainment involve examination of the proportions of DCO (death certificate only) registrations in South East England, and comparisons of cancer registration data with data from other clinical sources have been undertaken in the Netherlands and in Scotland. The latter studies demonstrated the extent of errors and incompleteness in cancer registry records, as well as shortfalls in ascertainment, and led to the conclusion (for example, in a study of registration of intracranial tumours) that cancer registries should use

a range of additional data sources to overcome these problems.¹¹

Relatively low ascertainment values in cancer registries could result in their data being biased in important respects. This is a potentially serious problem, which could lead to errors of interpretation. Routine linkage of data from a range of data sources provides a means to mitigate this. Such linkage is best done locally, because readier access to the data than is possible at the regional level reduces the likelihood of missed matches. We suggest that district registers, based on histopathology but augmented by other data sources, are a feasible way of providing accurate information both for immediate use as well as for regional and national purposes. This suggests a way forward, where the development of district registers could be linked to that of comprehensive information systems within local cancer units. This would have the advantage of improving the quality of data available to the register through access to operational systems collecting data in real time rather than retrospectively, while at the same time enhancing the quality of data in those same operational systems by creating additional mechanisms for data validation.

This work was funded by research grants from the Cornwall and Isles of Scilly Health Authority and Glaxo.

- 1 Expert Advisory Group on Cancer. *A policy framework for commissioning cancer services: a report to the chief medical officer (Calman-Hine report)*. London: Department of Health, 1995.
- 2 OPCS. *Review of the national cancer registration system (Alberman report)*. London: HMSO, 1990.
- 3 Codling BW, Pheby D, Hagen DL, et al. Cancer registration by linking pathology and district PAS data. *Br J Cancer* 1990;**62**:271-4.
- 4 Robles SC, Marrett LD, Clarke EA, et al. An application of capture-recapture methods to the estimation of completeness of cancer registration. *J Clin Epidemiol* 1988;**41**:495-501.
- 5 Brewster D, Crichton J, Muir C. *Accuracy of 1990 cancer registration data in Scotland*. Edinburgh: Common Services Agency, 1994.
- 6 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**:635-9.
- 7 Moss SM, Smith JAE, Nicholas DS. The quality of histopathology data in a computerised cancer information system: implications for future audit of care. *Public Health* 1997;**111**:101-6.
- 8 Quirke P, Williams GT. *Minimum dataset for colorectal cancer histopathology reports*. Royal College of Pathologists (<http://www.rcpath.org/activities/publications/ccancer.html>).
- 9 Comber H. Some patients with colorectal cancer may have been missed by Northern Ireland registry [letter]. *BMJ* 1999;**319**:385.
- 10 Seddon DJ, Williams EMI. Data quality in population-based cancer registration: an assessment of the Merseyside and Cheshire cancer registry. *Br J Cancer* 1997;**76**:667-74.
- 11 Counsell CE, Collie DA, Grant R. Limitations of using a cancer registry to identify incident primary intracranial tumours. *J Neurol Neurosurg Psychiatry* 1997;**63**:94-7.

WHAT'S YOUR VIEW?

***Pathology Interactive*: useful or useless?**

To make *Pathology Interactive* a more useful tool to help you in your CPD

... we need your views

Please take time to complete a self completion questionnaire and fax it back to us

Questionnaire available for printing out on:

Pathology Interactive CD Vol 1 issue 3 (Dec 1999)

Pathology Interactive CD Vol 2 2000

Journal of Clinical Pathology website - www.jclinpath.com

Fax completed forms to Caroline Clark, Editorial Assistant, *Journal of Clinical Pathology* (+44 (0)20 7383 6668).

Your time and help are appreciated.

***Pathology Interactive* 2000 offers up to three CPD credits per article; later issues feature case study and picture quiz formats**