Eur J Cancer Clin Oncol, Vol. 25, No. 10, pp. 1467–1473, 1989.
Printed in Great Britain

0277–5379/89$3.00 + 0.00

Making Cancer Statistics more Informative: Measures of the Quality of Recorded Diagnosis in a Population-based Registry

PAUL B.S. SILCOCKS,*† HELEN THORNTON-JONES‡ and RICHARD G. SKEET‡§

**Department of Clinical Epidemiology & Social Medicine, St. George's Medical School, Tooting, London SW17 0RE, U.K. and ‡Thames Cancer Registry, Sutton, Surrey SM2 5PY, U.K.*

Abstract—*The quality of the recorded diagnosis is a major limit to the usefulness of Cancer Registry statistics that is easily overlooked by users of the data. With data from a large population-based cancer registry as an example, we demonstrate how Registry statistics could be improved by wider use of three simple indices, namely* (1) *the proportion histologically verified (adjusted for age),* (2) *the proportion of verified cases with an uninformative diagnosis, and* (3) *the proportion of cases that are staged. We believe that greater awareness of the deficiencies of Cancer Registry statistics will lead to a more critical interpretation of them, and help stimulate efforts to rectify matters.*

INTRODUCTION

'THE broad purpose of cancer registration is to help assess and control the impact of malignancies on the community' [1]. In order to achieve these goals, population-based registries with geographically defined catchment populations operate throughout the world. In England and Wales these are usually based on a Health Region. Thames Cancer Registry (TCR), the registry responsible for the four Thames Health Regions, covers a population of about 14 million and registers some 55,000 cases per year. In terms of population covered it is currently the third largest registry in the world, although for the period covered by this paper the Registry was responsible only for the populations of the two South Thames Health Regions (population ca. 6.5 million).

At present we are compiling a report based on 25 years' worth of data stored in the Registry. Part of the work includes an assessment of the validity of the diagnoses as recorded in the Registry database. Establishing a diagnosis is the clinician's responsibility: physicians and surgeons decide whether to verify a tumour histologically, pathologists are responsible for specifying histological type and all three specialists play a part in the staging of cancer.

Accepted 9 June 1989.
†Present address and address for correspondence: ICRF Cancer Epidemiology Unit, Gibson Building, Radcliffe Infirmary, Oxford OX2 6HE, U.K.
§Present address: Herefordshire Health Authority, Victoria House, Hereford HR4 0AN, U.K.

Registries can indicate the overall validity of the diagnoses that have been recorded in a number of ways, such as the proportion of cases that are microscopically verified or the proportion registered only on the basis of a death certificate [2]. While the proportion of 'death certificate only' (DCO) registrations is indirectly a measure of the overall validity of diagnosis (in that no DCO registrations can be proved to have been verified), it is more an indicator of the Registry's timeliness in finding cases and of the extent to which medical records can be linked retrospectively, rather than of the quality of diagnosis *per se*. Because DCO registrations cannot be proved to have been histologically verified, if the proportion of DCO registrations is high, a truer index of the extent to which clinicians have established the diagnosis would be the proportion histologically verified having *excluded* DCO registrations. The range of performance indicators chosen by a Registry will of course depend on the uses to which Registry data are put. Besides indicators of the validity of diagnosis, other kinds of index are needed to monitor ascertainment and follow-up. The wider the variety of indicators available, however, the more readily different kinds of shortcomings in Registry data can be detected.

In this paper we emphasize the use of the proportion of cases histologically verified (adjusted for age) and we suggest two other indicators that can be used to monitor the diagnostic quality of Registry data, illustrating them with some of the results of our experience at TCR (although the issues involved

are common to all Cancer Registries). We also draw attention to ways in which the quality of recorded diagnoses can vary between sites and over time. This variation may not be appreciated by the naive user of Registry data nor even by the Registry itself if the relevant data are not recorded. While our data are not representative in the way that a *random* sample of cancer patients would be, they represent nevertheless a significant contribution to National Statistics because of the size of the population covered in the two South Thames Health Regions—larger than that of Denmark for example.

MATERIALS AND METHODS

General

In addition to the essential details on age, sex and date of diagnosis, since its inception in 1958 the Registry has recorded whether the diagnosis was based on a histological, cytological or clinical opinion. This is possible because the Registry uses field workers who obtain diagnosis-related data directly from the patients' case notes whenever they can, rather than from secondary sources such as hospital information systems (which are, however, used to *identify* cases).

Histological and cytological diagnoses are recorded at the registry using the *International Classification of Diseases for Oncology (ICD-O)* [3], the codes of which are identical with the tumour morphology codes of the Systematized Nomenclature of Medicine (SNOMED) [4]. At TCR, pathology *slides* are not reviewed, but wherever possible the diagnosis is coded from a copy or transcript of the original pathology report. This helps to avoid the inconsistencies in coding practice that would arise if the pathologist's own code were used and is analogous to the central coding of death certificates at the Office of Population Censuses & Surveys.

Before 1982 if the treating consultant had not specifically recorded the clinical TNM stage [5] in the case notes, clerks employed by the registry attempted to infer the T, N and M categories using whatever data were available. This was unsatisfactory and since 1982 a simpler procedure has been adopted in which the clerk notes the presence of local extension, lymph node involvement or distant metastases as recorded in the case notes, coding clinical opinions separately from surgical/pathological assessments. If a clinical TNM stage or other stage has also been specified in the case notes or if the TNM categories have been stated explicitly, these are recorded as well. While it is not possible from registry records to distinguish between surgically staged cases with and without histology, we consider that surgical/pathological staging will inherently be more accurate than a clinical opinion alone and so in this paper we have examined only surgical/pathological staging.

Quality of diagnosis indices

From our experience at TCR we decided that the quality of recorded diagnoses could best be summarized in terms of: verification that cancer was present, specification of its type and whether the tumour was staged.

Verification. The proportion of cases in which the diagnosis was histologically verified was examined for five time periods (1961–64, 1965–69, 1970–74, 1975–79 and 1980–84). The proportion of DCO registrations was also obtained for the last time period, in order to assess the extent to which this might affect the proportion verified. The proportions verified during each time period were indirectly age-standardized using the age-specific proportions verified for all sites (persons) in 1984 at TCR as a standard.

Specification. For each site the tumour morphology codes recorded during 1980–84 were grouped according to a registry-based classification [6]. The aim of the classification is to group *ICD-O* codes into fairly broad hierarchical categories in order to avoid the problem of diagnostic transfer and to ensure that none of the variety of morphology codes used in practice are wasted. The period 1980–84 was chosen because only diagnoses recorded since 1979 will have been coded directly to *ICD-O*. We considered that the following 3 digit *ICD-O* morphology codes were likely to reflect a vague and therefore inadequately specified diagnosis: 800 (neoplasms, not otherwise specified), 801–804 (epithelial neoplasms, not otherwise specified having excluded small and oat cell carcinomas—codes 8042, 8043), and 880 (sarcoma, not otherwise specified).

Staging. The proportion of cases staged using the TNM or, where appropriate, other standard method was also recorded for each site and compared with the proportion staged using the simplified method adopted at TCR. We counted a case as staged if T, N and M categories or the stage derived from them were recorded.

Sites chosen

We chose to illustrate these indices using cancers of the lung, stomach, breast, colon and rectum because of their frequency. In addition we included pancreas and brain because we predicted that advances in imaging technology would have affected the quality of diagnosis for these sites.

Statistical methods

Confidence limits for the age-standardized proportion verified were calculated using the square-root transformation on the observed number veri-

Table 1. *Percentage of cases histologically verified by age and site* 1980–84 (*to nearest* %)—*persons*

Site	Age band (years) <55	55–64	65–74	75+	χ^2 (1 df) (Trend)
Lung	72	63	49	24	3376.6
(*n*)	(1810)	(5665)	(10,345)	(8321)	
Stomach	81	77	69	46	555.6
(*n*)	(490)	(1086)	(2521)	(3214)	
Colon	87	82	77	58	612.4
(*n*)	(917)	(1663)	(3199)	(4633)	
Rectum	91	89	85	69	256.1
(*n*)	(544)	(1109)	(2013)	(2535)	
Female breast	89	83	76	50	1664.2
(*n*)	(4351)	(3710	(4181)	(4216)	
Brain	73	66	50	23	239.0
(*n*)	(746)	(501)	(526)	(215)	
Pancreas	59	47	35	19	348.9
(*n*)	(307)	(761)	(1506)	(1782)	

Table 2. *Percentage of cases histologically verified adjusted for DCO registrations by age and site* 1980–84 (*to nearest* %)—*persons*

Site	Age band (years) <55	55–64	65–74	75+
Lung	75	67	53	27
Stomach	86	82	74	52
Colon	91	87	82	65
Rectum	94	92	89	74
Female breast	91	86	79	56
Brain	76	70	54	26
Pancreas	64	52	39	21

fied (assuming it to have a Poisson distribution). The arcsin $\sqrt{p}$ transformation [7] was used for binomial proportions; the standard error of the relative survival rate (which standardizes for different ages at diagnosis) was found using (1/*Pe*) × Greenwood's estimate [8] of the standard error of the survival curve, where *Pe* is the expected proportion of survivors. Chi-square tests were constructed (where appropriate) as described by Fleiss [9].

RESULTS

Verification

Table 1 shows that the proportion of cases verified histologically varies markedly by age and site with malignancies in older persons being much less likely to be histologically verified. The results of allowing for DCO registrations are shown in Table 2. Tables 3 and 4 show three patterns with time in the age-standardized proportion verified: a steady upward trend in the proportion verified (as for tumours of the colon and rectum), little change over time (as with tumours of the lung) and an inverted-U shaped trend (as seen with tumours of the brain in both sexes). For brain tumours the peak period for histological verification was 1970–74 and for male cases of pancreatic cancer there is also a hint of a peak in 1975–79. Sex differences exist too: as a general rule males are slightly more likely to have histologically verified tumours than females.

Specification

For histologically verified cases, the proportion of 'inadequately specified' diagnoses is given by site in Table 5. The site with the highest proportion of such diagnoses was breast (30%) and the lowest was brain (2%).

Staging

Few cases were staged according to the TNM system (Table 6). The simpler scheme adopted by the Registry (Table 7) gives prognostic information (Table 8) as does the TNM system, but can be applied to a much higher proportion of cases. It is not applicable to primary brain tumours, however, because these rarely spread beyond the organ of origin.

DISCUSSION

Verification

For deep seated organs, at least, histological verification of diagnosis increases one's confidence that a neoplasm existed and that it arose at the stated site [2]. Nevertheless, for sites of major importance only about half of the registered cases may be histologically verified. While this may partly reflect the availability of non-invasive methods of diagnosis and the clinicians' wish to avoid needless investigation, even with accessible and treatable sites such as breast and large bowel the proportion

Table 3. Age-standardized percentage of cases histologically verified by period of diagnosis and site (males)—to nearest % (95% confidence limits)

Sex	Site	1960–64	1965–69	1970–74	1975–79	1980–84	χ^2 (df = 1) Trend	χ^2 (df = 3) Non-linearity
Males	Lung	47 (46–48)	47 (46–48)	45 (44–46)	46 (45–47)	47 (46–48)	0.4	13.2
	Stomach	47 (45–49)	53 (51–55)	59 (57–61)	65 (63–67)	67 (65–69)	224.7	4.6
	Colon	65 (62–68)	69 (66–72)	72 (69–75)	74 (71–77)	75 (72–78)	30.7	1.8
	Rectum	77 (74–81)	80 (76–83)	83 (80–86)	85 (82–88)	85 (82–89)	15.6	1.43
	Brain	59 (55–63)	65 (60–70)	69 (64–74)	59 (53–65)	55 (51–59)	4.1	19.1
	Pancreas	27 (24–30)	33 (30–36)	34 (31–37)	35 (33–37)	33 (31–35)	11.5	11.0

Critical values of χ^2_1: 3.841 ($P = 0.05$), 6.635 ($P = 0.01$), 10.827 ($P = 0.001$); χ^2_3: 7.815 ($P = 0.05$), 11.341 ($P = 0.01$), 16.268 ($P = 0.001$).

Table 4. Age-standardized percentage of cases histologically verified by period of diagnosis and site (females)—to nearest % (95% confidence limits)

Sex	Site	1960–64	1965–69	1970–74	1975–79	1980–84	χ^2 (df = 1) Trend	χ^2 (df = 3) Non-linearity
Females	Lung	45 (42–47)	45 (43–47)	45 (43–47)	45 (43–47)	46 (44–48)	0.6	0.6
	Stomach	44 (42–46)	50 (48–52)	52 (50–54)	57 (54–60)	60 (57–63)	89.1	2.0
	Colon	67 (65–69)	70 (68–72)	71 (69–73)	72 (70–74)	74 (72–76)	17.8	0.9
	Rectum	75 (72–78)	77 (74–80)	79 (76–82)	81 (78–84)	81 (78–84)	8.9	0.6
	Brain	55 (50–60)	62 (57–67)	62 (57–67)	57 (52–62)	52 (47–57)	2.3	10.7
	Pancreas	26 (23–29)	34 (31–37)	32 (29–35)	33 (30–36)	34 (31–37)	12.7	12.0
	Breast	68 (67–69)	71 (70–72)	70 (69–71)	72 (71–73)	71 (70–72)	10.5	9.7

Critical values of χ^2_1: 3.841 ($P = 0.05$), 6.635 ($P = 0.01$), 10.827 ($P = 0.001$); χ^2_3: 7.815 ($P = 0.05$), 11.341 ($P = 0.01$), 16.268 ($P = 0.001$).

verified is still well short of 100% despite recognition of the *clinical* importance of obtaining a histological diagnosis [10]. For the sites studied, adjustment for DCO registrations produced little effect on the percentages histologically verified which may therefore be taken to reflect directly the willingness of clinicians to establish the diagnosis. We therefore omitted the adjustment from further analyses. Because the site, time period and sex under consideration may affect the age-distribution of cancer patients it is advisable to allow for the effect of age when comparisons are made. A simple way to do this is by age-standardizing the proportions verified. This information is now available for international data [11, 12] and is one of the indices whose wider use by Registries has been encouraged [11]. However, further insights can be gained by examining time trends in histological verification. For example, they may also reflect changes in other diagnostic methods, thus the pattterns seen in brain and pancreatic tumours are consistent with the introduction of brain scanning in the early 1970s

Table 5. Histologically verified cases with an inadequately specified morphological diagnosis, by site 1980–84: to nearest %

Site No. verified	Percentage with inadequately specified morphology (95% confidence limits)
Lung (n = 11,926)	27 (26.4 – 28.1)
Stomach (n = 4444)	17 (15.7 – 18.0)
Colon (n = 7343)	7 (6.7 = 8.0)
Rectum (n = 4946)	5 (4.7 = 6.0)
Brain (n = 1186)	2 (1.2 = 2.8)
Pancreas (n = 1391)	23 (21.1 – 25.6)
Female breast (n = 12,246)	30 (29.4 = 31.1)

Table 6. Percentage of cases staged by TNM, other standard system and TCR system 1982–84 by site

Site	Percentage staged by TNM	Percentage staged by Other	Percentage with TCR surgical/pathological staging
Lung (n = 15,608)	3	0.1	34
Stomach (n = 4277)	2	0.5 (Dukes)	56
Colon (n = 6363)	0.6	23 (Dukes)	71
Rectum (n = 3719)	0.9	72 (Dukes)	75
Brain	← not applicable →		
Pancreas (n = 2595)	0	0	44
Female breast (n = 10,192)	13	5	65

Table 7. Simplified staging for Cancer Registry

System at TCR	Condensed version (as used in this paper)
1 Localized to organ of origin	1 Localized to organ of origin
2 Direct extension from organ of origin, into adjacent tissues	2 Direct extension from organ of origin into adjacent tissues
3 Localized + nodes	3 Nodes involved
4 Direct extension + regional nodes	
5 Localized + distant metastases	4 Distant metastases
6 Direct extension + distant metastases	
7 Localized + nodes + distant metastases	
8 Direct extension + nodes + distant metastases	
9 Metastases, status of tumour at primary site not known	

*Table 8. Percentage of 3-year relative survival by surgical/pathological TCR stage by site–persons registered in 1982–83 (standard error)**

Site	TCR stage Localized	Direct extension only	Lymph nodes involved	Distant metastases	Not staged
Lung	19 (1.7)	17 (1.4)	18 (1.8)	4 (0.8)	6 (0.4)
Stomach	28 (3.2)	11 (1.8)	17 (1.9)	3 (0.9)	8 (1.2)
Colon	71 (2.2)	54 (2.3)	37 (2.2)	11 (1.3)	15 (1.8)
Rectum	65 (2.5)	49 (2.9)	37 (2.8)	9 (1.7)	18 (2.3)
Brain	← not applicable →				
Pancreas	10 (2.3)	6 (2.2)	7 (3.5)	4 (1.1)	4 (1.0)
Female breast	92 (0.8)	83 (2.1)	75 (1.3)	41 (4.1)	55 (1.4)

*When survival is poor, asymmetric confidence limits may be required. These can be found by estimating the binomial sample size N that would give the same standard error for the same proportion surviving. Exact or approximate methods for calculating binomial confidence limits are then applied [12].

[13] and the introduction of body scanners near the end of the decade.

Specification

While the fact of cancer may be fairly reliably determined by modern non-invasive methods, such techniques must still be validated and are unlikely to give information on histological type. But even when a case is histologically verified the diagnosis recorded at the registry may be vague and the proportion of such diagnoses which we regard as inadequately specified (as defined in the Materials and Methods section of this paper) varies considerably between sites for reasons that are as yet obscure. For example the proportion of 'inadequately specified' cancers is four times greater for the pancreas than for the rectum, although in both sites most tumours are adenocarcinomas. It might be thought that because a high percentage of pancreatic cancers is inoperable, there may be relatively little tissue available and little incentive to give a precise diagnosis. However, the same ratio is found in treated cases. Furthermore, women with treated breast tumours are nearly six times more likely than women with treated rectal cancers to have an 'inadequately specified' diagnosis, although for both sites the prognosis is relatively good and treatment normally provides a reasonable quantity of tissue on which to make a diagnosis. At present, however, the TCR database does not indicate the type or size of specimen on which histological diagnoses are based.

Our definition of 'inadequately specified' tumours is admittedly crude and may need to be refined for some sites. For example, although 98% of histologically verified brain tumours had what we termed an 'adequately specified' diagnosis, just over a third of these were 'glioma, malignant', which is clearly inadequate for many purposes. On the other hand, we may have been too severe in our definition of 'inadequately specified' tumours because it included 'undifferentiated epithelial neoplasms' (defined by us as 3 digit *ICD-O* codes 802–804, excluding small and oat cell carcinomas) and 'sarcomas NOS' (*ICD-O* code 880). Because undifferentiated tumours undoubtedly exist these terms might be regarded as more circumstantial than 'neoplasm not otherwise specified'. However, whether or not a tumour is classed as undifferentiated will depend on how hard the pathologist searches for evidence of differentiation, the availability of electron microscopy and immunohistochemistry, the size of the specimen and the extent to which it is representative of the tumour. Because the variability of these factors is unknown we deliberately chose to be conservative. Many of the above limitations could easily be rectified if a single site were under review, and we believe that a search for such uninformative histological terms can be a useful quality control activity. If histological type is relevant to prognosis or aetiology, it is worth recording properly. If it is not being recorded properly an active decision should be made whether to insist on an improvement or to cease recording this information.

Tumours that have not been histologically verified should receive a (3 digit) '999' morphology code. However, if the fact of histological verification is recorded independently of diagnosis (as it is at TCR) then a morphology code can be used to take account of a clinical impression (for example that a skin malignancy is a melanoma), while maintaining compatability with 'standard' *ICD-O* practice. Thus even tumours for which there is no record of histological or cytological verification *may* be given quite circumstantial sounding diagnoses (4.6% of non-verified lung cancers for example) but such diagnoses should not in general be taken at their face value.

Staging

This provides an index of the extent to which a tumour has spread. Cancer registries need to record stage so that comparisons of survival between treatment centres or over periods of time are made between similar types of patient. In addition, if stage is associated with age or sex, there may be implications for the deployment of preventive services. Although staging rules have existed for many years only a small proportion of cases have stage either explicitly stated or implied by clinicians in the case notes. With a simple staging system it is clear that at least some information of prognostic value can be recorded even when formal staging has not been done. However, until such a system becomes widely used, most cancer registration data will either be presented regardless of stage (thus losing insights into prognosis) or stage-related tables will be based on a small, unrepresentative subgroup of patients.

Future prospects

In the future, cancer registration may be improved by direct communication between laboratory and Registry computers, although this will not affect the quality of diagnosis in cases that are not verified nor avoid unhelpful inadequately specified diagnoses. In addition it is our impression that laboratory-based notifications are less accurate as to site. Review of actual histology slides would help improve the recording of cancer type, but would require extra money and facilities. Less expensive changes should be implemented first. Wider use of age-adjusted indicators of verification, and simple indices of the extent to which cancers are histologically specified and staged can improve Registry data by highlighting variations in the quality of the recorded diagnosis. Greater awareness of such variation will stimulate efforts to understand its origins and to rectify it: thus it may be helpful to record the type and size of the specimen on which a diagnosis was based and a separate axis may need to be introduced for recording immunohistochemical investigations, as is rapidly becoming necessary for malignant lymphomas.

REFERENCES

1. MacLennan R, Muir C, Steinitz R, Winkler A. *Cancer Registration and its Techniques.* IARC Publications No. 21. Lyon, International Agency for Cancer Research, 1978.
2. Reliability of registrations. In: Waterhouse J, Muir C, Correa P, Powell J. *Cancer Incidence in Five Continents III.* Lyon, International Agency for Cancer Research, 1976, Ch. VI, 45–50.
3. World Health Organization. *International Classification of Diseases for Oncology (ICD-O).* Geneva, WHO, 1976.
4. College of American Pathologists. *SNOMED. Coding manual; alphabetic index; numerical index.* Skokie, Illinois, College of American Pathologists, 1979.
5. Hermanek P, Sobin LH, eds. *TNM Classification of Malignant Tumours,* 4th edn. Berlin, Springer, 1987.
6. Silcocks PB, Thornton-Jones H. A classification scheme for childhood cancer (letter). *Int J Cancer* 1988, **42**, 642–644.
7. Armitage P, Berry G. *Statistical Methods in Medical Research,* 2nd edn. Oxford, Blackwells Scientific Publications, 1987.
8. Greenwood M. *Reports on Public Health & Medical Subjects,* No. 33, Appendix 1. London, HMSO, 1926.
9. Fleiss JL. *Statistical Methods for Rates and Proportions,* 2nd edn. New York, John Wiley, 1981.
10. Buck N, Devlin HB, Lunn JN. *The Report of a Confidential Enquiry into Perioperative Deaths.* London, The King's Fund, 1988.
11. Comparability of data and reliability of registrations. In: Waterhouse J, Muir C, Shanmugaratnam K, Powell J, eds. *Cancer Incidence in Five Continents IV.* IARC Scientific Publications No. 62. Lyon, International Agency for Cancer Research, 1982, Ch. 6, 55–67.
12. Comparability of data and relability of registrations. In: Waterhouse J, Muir C, Mack T, Powell J, Whelan S, eds. *Cancer Incidence in Five Continents V.* IARC Scientific Publications No. 88. Lyon, International Agency for Cancer Research, 1987, Ch. 7, 45–61.
13. Hounsfield GN. Computerized transverse-axial scanning (tomography) part 1: description of system. *Br J Radiol* 1973, **46**, 1016–1022.
14. Rothman KJ. Estimation of confidence limits for the cumulative probability of survival in life table analysis. *J Chron Dis* 1978, **31**, 557–560.