

COLON CANCER

Accuracy of reporting of family history of colorectal cancer

R J Mitchell, D Brewster, H Campbell, M E M Porteous, A H Wyllie, C C Bird, M G Dunlop

Gut 2004;53:291–295. doi: 10.1136/gut.2003.027896

See end of article for authors' affiliations

Correspondence to: Professor M G Dunlop, Academic Coloproctology, University of Edinburgh, Edinburgh EH4 2XU, UK; Malcolm.Dunlop@hgu.mrc.ac.uk

Accepted for publication 30 September 2003

Background and aims: Family history is used extensively to estimate the risk of colorectal cancer but there is considerable potential for recall bias and inaccuracy. Hence we systematically assessed the accuracy of family history reported at interview compared with actual cancer experience in relatives.

Methods: Using face to face interviews, we recorded family history from 199 colorectal cancer cases and 133 community controls, totalling 5637 first and second degree relatives (FDRs/SDRs). We linked computerised cancer registry data to interview information to determine the accuracy of family history reporting.

Results: Cases substantially underreported colorectal cancer arising both in FDRs (sensitivity 0.566 (95% confidence interval (CI) 0.433, 0.690); specificity 0.990 (95% CI 0.983, 0.994)) and SDRs (sensitivity 0.271 (95% CI 0.166, 0.410); specificity 0.996 (95% CI 0.992, 0.998)). There was no observable difference in accuracy of reporting family history between case and control interviewees. Control subjects similarly underreported colorectal cancer in FDRs (sensitivity 0.529 (95% CI 0.310, 0.738); specificity 0.995 (95% CI 0.989, 0.998)) and SDRs (sensitivity 0.333 (95% CI 0.192, 0.512); specificity 0.995 (95% CI 0.991, 0.995)). To determine practical implications of inaccurate family history, we applied family history criteria before and after record linkage. Only two of five families reported at interview to meet surveillance criteria did so after validation, whereas only two of six families that actually merited surveillance were identified by interview.

Conclusions: This study has quantified the inaccuracy of interview in identifying people at risk of colorectal cancer due to a family history. Colorectal cancer was substantially underreported and so family history information should be interpreted with caution. These findings have considerable relevance to identifying patients who merit surveillance colonoscopy and to epidemiological studies.

People who have relatives affected by colorectal cancer have an increased personal risk of the disease compared with the general population. The degree of personal risk relates to the extent of family history and age of onset of affected relatives.¹ Thus family history is used in the clinical setting to inform decisions regarding the use of colonoscopic surveillance. Because of the increasing awareness of the genetic contribution to colorectal cancer, in the UK, elsewhere in Europe, and in the USA there has been a rapid increase in colonoscopy workload where family history is the primary concern. Guidelines based on degree of family history have been devised to determine when surveillance should be recommended.^{2–3} This empiric approach inherently places considerable importance on the accuracy of family history information. Accuracy is also an important consideration in the context of the epidemiological studies that inform the guidelines for offering surveillance. In both situations, information on family history is usually gathered by interview with a family member. This approach is potentially subject to inaccuracy on the part of the interviewee. Underreporting of family history has been observed in previous studies^{4–5} and there is evidence that systematic recall bias may arise from the fact that people with raised awareness of a particular cancer may be more likely to report a positive family history.⁶ Furthermore, the social stigma associated with bowel cancer may mean that this condition is discussed less readily within families, and this factor could particularly affect reporting of family history.

Accuracy of reporting cancer in the family has been addressed in previous studies of people referred to genetics departments because of a cancer family history,^{7–9} people with a personal history of cancer,^{5, 6, 10–14} or close relatives of cancer cases.^{4, 15, 16} However, only a few studies have related specifically to colorectal cancer cases^{4, 5, 11, 12, 14} or to commu-

nity based consultands who have not been referred to a genetics clinic.^{5, 10} Another limitation of the published literature is that validation of the interviewee's report is often only attempted for relatives reported to have had cancer. In such studies no information can be obtained regarding the sensitivity, specificity, or negative predictive value of reports, and the question of underreporting cannot be addressed.

In this study, information obtained at interview from colorectal cancer cases and community controls was linked systematically to Scottish Cancer Registry data in order to investigate the true accuracy of reporting of a family history of colorectal cancer. We determined the cancer experience of 5637 relatives, irrespective of the reporting of cancer by the interviewee, and so we were able to determine overall accuracy, including underreporting of cancer in relatives. We also evaluated the effect of any inaccuracies on clinical interpretation of family history with respect to recommending surveillance colonoscopy. The findings have considerable relevance to the methods used to validate family history and also have practical implications for surveillance guidelines.

METHODS

A genetics nurse conducted face to face interviews with cases and controls to obtain their reported family history. A total of 199 consecutive colorectal cancer cases were ascertained from Edinburgh Royal Infirmary, Western General Hospital, Edinburgh, and St Johns Hospital, Livingston. For community controls, our initial strategy was to recruit spouses of

Abbreviations: FDR, first degree relatives; SDR, second degree relatives; ISD, Information and Statistics Division

Table 1 Consultant knowledge of all cancer types in relatives

Interviewee group	Relative group	No of relatives	No (%) for whom interviewee could supply health information	No of relatives with confirmed cancer*	No (%) of affected relatives in which cancer was reported	Total No of cancers†	No (%) of cancers accurately reported
Cases	FDR	1322	1250 (95%)	215	152 (71%)	240	106 (44%)
Cases	SDR	1968	713 (36%)	274	84 (31%)	293	42 (14%)
Controls	FDR	1037	991 (96%)	113	76 (67%)	124	51 (41%)
Controls	SDR	1310	671 (51%)	189	77 (41%)	202	36 (18%)

*This column refers to the total number of relatives in a particular group found by ISD linkage to have had cancer.

†This column describes the total number of primary cancers occurring in relatives, including multiple primary cancers.
FDR, first degree relatives; SDR, second degree relatives.

cases. However, this approach proved impractical, and only 25 controls were identified by this means. A further 108 age and sex matched controls were ascertained from general practice lists in North West Edinburgh. Details of all first and second degree relatives (FDRs/SDRs), as reported by the interviewee, were recorded in a structured proforma. A comprehensive manual search of records of births, deaths, and marriages held at the General Register Office for Scotland was performed, in order to verify, correct, and extend pedigree information reported at interview in preparation for record linkage.

Data for all relatives were systematically linked to Scottish Cancer Registry data held by the Information and Statistics Division (ISD) of the Scottish Executive. The Scottish Record Linkage System links all records relating to hospital discharge, cancer registration, and cause of death for each individual, and represents a comprehensive resource for identifying cancer incidence in a given population group. Using techniques based on the principles of "probability matching" developed by Newcombe,¹⁷ such records are linked via patient specific identifying information with a false positive rate of less than 1%.^{18–20} Our own internal assessment of colorectal cancer ascertainment is that the false negative rate is also of this order. The same methodology can be applied to linking research data containing personal identifiers with the health information held by ISD. Surname, forename, sex, date of birth, and postcode are commonly used to match records, and our data set contained all but the latter of these. Record linkage served not only to validate reports of cancer but also to identify previously unidentified cases. Confidence intervals (CI) were calculated using an approximation based on inverting an appropriate score test statistic,²¹ which compares favourably with exact methods for our data.²²

Ethics approval for the recruitment and interview of patients and controls was granted by the Lothian Local Research Ethics Committee. The record linkage process was subject to approval from the Privacy Advisory Committee responsible for advising ISD on the release of patient identifiable data. All linked data remained at ISD throughout the analyses to ensure confidentiality.

RESULTS

Mean age of the 199 colorectal cases at the time of interview was 64.0 years. There were 86 females and 113 males, who had a total of 3290 relatives included in the database. One hundred and ten relatives were reported to be resident outside Scotland, and the nurse constructing the pedigrees classified a further 251 as "untraceable". Mean age of the 133 controls was 64.2 years at the time of interview. There were 60 females and 73 males with a total of 2347 relatives. In all there were 107 relatives who were reported to be resident outside Scotland and 91 were deemed to be untraceable. Individuals who have neither died nor developed cancer will not be matched through record linkage, and so it is impossible to distinguish these individuals from those who

cannot be traced. Hence all 3290 relatives of cases and 2347 relatives of controls were included in the subsequent record linkage and analysis, regardless of apparent "traceability".

Knowledge of family members' health and occurrence of all types of cancer

Interviewees were asked to state their knowledge as to whether a given relative was alive, and regarding the medical history of relatives, including any history of cancer. The proportion of relatives for which the interviewees were able to provide any health related information is shown in table 1. Table 1 also details the responses given by interviewees for all relatives found to have any type of cancer by linking with central records.

In the majority of instances where a cancer was not correctly reported, the interviewee either had no knowledge of the health of the relative in question or was unaware that they had developed any type of cancer. However, in some cases a cancer was reported but the site was incorrect or unknown. An indication of the extent to which this occurred is provided by the sixth column in table 1, which states the proportion of affected relatives reported to have had any form of cancer.

Reporting of colorectal cancer cases

There were a total of 148 confirmed cases of colorectal cancer in FDRs or SDRs, of which 62 were reported correctly by the interviewee. Mean age at onset of cases that were correctly reported was 63.3 years (95% CI 60.5, 66.1), a value significantly different from the mean age of 70.2 years (95% CI 67.8, 72.5) for cases that were not correctly reported. This observation is not unexpected as cancer affecting more elderly relatives is less likely to be discussed within families. The suggestion that early onset cases are more likely to be reported accurately at interview is of clinical interest as such cases are more significant in terms of indicating increased genetic risk. A separate trend towards more accurate reporting in recent years was evident, although not statistically significant. Summary statistics associated with the accuracy of reporting of colorectal cancer in relatives are presented in table 2.

The data in table 2 demonstrate substantial underreporting of colorectal cancer in relatives. In both cases and controls, sensitivity of reporting in FDRs is approximately 50–60%, implying that a large proportion of cancers in FDRs go unreported. The poor sensitivity of reporting is even more striking in SDRs, with the majority of cases in SDRs of cases and controls not being reported at interview. The very high estimates of specificity and negative predictive value primarily reflect the fact that in absolute terms colorectal cancer affects only a small proportion of the population. However, even small effects on these parameters may have important implications for genetic risk assessment and resource allocation. For all relative groups, estimates of positive predictive value were in the range 60–70%, indicating that approximately one third of reports of individual colorectal

Table 2 Summary statistics for reporting of colorectal cancer

Group	Relative group	No of cases reported at interview	No of cases found by record linkage	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)
Cases	FDR (n = 1322)	43	53	0.566 (0.433, 0.690)	0.990 (0.983, 0.994)	0.698 (0.549, 0.814)	0.982 (0.973, 0.988)
Cases	SDR (n = 1968)	21	48	0.271 (0.166, 0.410)	0.996 (0.992, 0.998)	0.619 (0.409, 0.792)	0.982 (0.975, 0.987)
Controls	FDR (n = 1037)	14	17	0.529 (0.310, 0.738)	0.995 (0.989, 0.998)	0.643 (0.388, 0.837)	0.992 (0.985, 0.996)
Controls	SDR (n = 1310)	16	30	0.333 (0.192, 0.512)	0.995 (0.991, 0.995)	0.667 (0.417, 0.848)	0.985 (0.976, 0.990)

Sensitivity, the proportion of colorectal cancer cases identified by record linkage that were reported at interview; Specificity, the proportion of relatives confirmed by record linkage to be unaffected that were reported to be colorectal cancer free at interview; Positive predictive value, the proportion of reported colorectal cancer cases that were confirmed by record linkage; Negative predictive value, the proportion of individuals reported to be free of colorectal cancer at interview that were confirmed as unaffected by record linkage.

FDR, first degree relatives; SDR, second degree relatives; 95% CI, 95% confidence interval.

cancer cases are not confirmed using cancer registry data. The sensitivity of reporting of colorectal cancer compared with other common cancers is shown in table 3. As no differences were observed between cases and controls in terms of the accuracy of family history reporting, all consultands have been grouped together.

Estimates of sensitivity for colorectal cancer were broadly comparable with the other common cancer types listed in table 3, although numbers were small. However, it is noteworthy that breast cancer was more frequently reported than the other internal cancers in FDRs. This may reflect the more enigmatic presentation of visceral malignancy and the social stigma associated with bowel cancer in particular.

Practical implications of inaccurate or incomplete reporting of family history

From a clinical perspective it is important to determine the validity of interviewee reporting as a means of identifying families that are eligible for colonoscopic surveillance and/or genetic testing. Various guidelines exist to help determine the extent of family history that warrants such interventions, but for illustrative purposes we have applied family history criteria adopted by the British Society of Gastroenterology and the Association of Coloproctology of Great Britain and Ireland (two FDRs with colorectal cancer, or one FDR diagnosed under 45 years).³ Using these family history criteria, we identified a group of interviewees who merited colonoscopic surveillance. We then re-evaluated the risk categorisation of these individuals based on validated family history data following record linkage.

Again, cases and controls were considered together. In order to gauge the overall impact of inaccurate or incomplete reporting on surveillance recommendations, cases and controls were considered simply as consultands, rather than cases meriting postsurgical surveillance following their own personal history of colorectal cancer. At interview, five of the interviewees reported a family history that met criteria indicating a need for surveillance. However, only two of these five families were confirmed by record linkage to meet these criteria, giving an overall positive predictive value of 0.400 (95% CI 0.118, 0.769). In addition, four further consultands who did not report a family history of colorectal cancer fulfilling criteria actually did have such a family history based on record linkage data. Therefore, only two of six consultands who should have been recommended for surveillance were identified at interview, suggesting that the sensitivity of interview in terms identifying appropriate individuals for surveillance is 0.333 (95% CI 0.097, 0.700).

DISCUSSION

This study has quantified the accuracy of reported family history of cancer in two important groups of people—namely, those with colorectal cancer and those from the general population. Because we confirmed cases reported to have colorectal cancer and also identified cases that had not been reported by the interviewee, we have been able to systematically assess overall accuracy of reported family history of large bowel malignancy.

Using this approach we have determined the accuracy of reporting of colorectal cancer in a large data set comprising 332 interviewees and 5637 first and second degree relatives. We showed conclusively that substantial underreporting of cancer family history is evident in reports made at interview. In this study, the family history documentation was optimal as a trained genetics nurse conducted interviews during a lengthy consultation at the interviewee’s home. Reporting inaccuracies may be more extreme where family history is taken in a busy gastroenterology, surgical, or general practice clinic.

Table 3 Sensitivity of interview as a means of identifying familial cancer cases, by site

Site	Relative group	No of cases*	No of cases correctly reported	Sensitivity of interviewee report (95% CI)
Colorectal	FDR	70	39	0.557 (0.441, 0.668)
Colorectal	SDR	78	23	0.295 (0.205, 0.404)
Breast	FDR	28	21	0.750 (0.566, 0.873)
Breast	SDR	37	11	0.297 (0.175, 0.458)
Bronchus and lung	FDR	66	37	0.561 (0.441, 0.674)
Bronchus and lung	SDR	67	10	0.149 (0.083, 0.253)
Stomach	FDR	30	12	0.400 (0.246, 0.577)
Stomach	SDR	64	11	0.172 (0.099, 0.282)

*Where more than one primary cancer occurred at the same site, it was not possible to determine whether the interviewee was aware of both of these tumours. Therefore, where metachronous primary cancers occurred, only the first is considered.

FDR, first degree relatives; SDR, second degree relatives; 95% CI, 95% confidence interval.

A comparable approach to assessing accuracy of reporting of colorectal cancer, which includes identification of unreported cases as well as checking the accuracy of cases reported at interview, has been employed in one previous study.⁵ This study estimated the sensitivity of reporting a family history of colorectal cancer in FDRs as 0.65 (95% CI 0.39, 0.85) for colon cancer cases and 0.81 (95% CI 0.54, 0.95) for controls, and the authors concluded that subjects were able to accurately report family history.⁵ However, this previous study did not consider SDRs, and no information is provided regarding the total number of relatives involved. Furthermore, the focus of this paper was on validation of an epidemiological study. The observed values for sensitivity of reporting may be less acceptable for genetic risk assessment where the objective is to determine the need for clinical intervention, particularly given the wide confidence intervals.

In general, there is a distinct lack of quality data regarding the accuracy of reporting of family history of colorectal cancer at interview, and the impact of inaccuracy and under-reporting on genetic risk assessment has not been evaluated. The current study is thus highly relevant, particularly given the current increase in public demand for information on genetic risk.

We did not observe any difference in the accuracy of family history reporting in cases compared with controls. Similarly, age and sex of interviewee had no significant effect on accuracy. Clearly, the accuracy of reporting of family history by colorectal cancer cases is an important consideration as cancer occurrence is frequently the first point of contact with a particular family. This study addresses the hypothesis that individuals who have had colorectal cancer may be more likely than controls to provide false positive reports of the condition in their relatives. However, we found no evidence to support this hypothesis as there were 21 false positive reports among 199 cases compared with 11 false positive reports among 133 interviewed controls.

Table 1 shows that interviewees could provide no useful information for approximately half of all SDRs but did have some knowledge of the health status of all but approximately 5% of FDRs. This consistent disparity suggests that many instances in which cancer in SDRs goes unreported are due to lack of contact with relatives, rather than ignorance of diagnosis in a known family member. The observation that positive predictive value is similar in FDRs and SDRs lends further support to this notion. Clearly, one would expect that interviewees would have greater knowledge about FDRs, and would be more likely to receive and maintain knowledge of a cancer diagnosis from such close family. Disparity between FDRs and SDRs is evident throughout this study, and is consistent with findings from other published studies.⁷⁻¹⁰

There is some potential for bias within this study but we feel that the effect of such bias is minimal. The total

proportion of potential participants who declined to take part in the study, or did not respond to a letter of invitation, was less than 20%. False positive and false negative rates were low for the record linkage process that we used, emphasising the overall validity of our approach. Spouses of cases may be more aware of their own family history of colorectal cancer than the general population, although any such effect would only apply to a small proportion of control subjects. Some mismatching may have occurred, and a proportion of relatives, probably approximately 10%, may have been untraceable. This latter effect would theoretically lead to an underestimation of the positive predictive value. However, no cases and only one control subject reported colorectal cancer in a relative reported to live abroad or deemed to be untraceable, and consequently this effect will have little influence on the reported results.

The accuracy and completeness of cancer registry data itself is a crucial consideration for any study that uses such a resource to validate or confirm diagnoses. The Scottish Cancer Registry was initiated in 1958, and ascertainment was considered to be suboptimal prior to 1968. Although ascertainment of any registry is unlikely to reach 100%, methods of ascertainment have steadily improved since this time, and the Scottish Cancer Registry is considered to be reasonably complete in recent years and to compare favourably with other registries.²³ An evaluation of the accuracy of colorectal cancer registration data found that while misclassifications do occur at a low level, such data exhibit a high degree of accuracy.²⁴ Colorectal cancer cases occurring prior to the availability of an effective cancer registry were only identified by this study if this malignancy was recorded as a cause of death. Again, this is unlikely to introduce systematic bias, but may have resulted in a slight underestimation of the positive predictive value. Overall, therefore, we consider record linkage with the Scottish Cancer Registry to constitute a reliable and valid means of determining the actual cancer experience of our study subjects. The intermediate use of central records to confirm or correct reported information and to extend knowledge of pedigrees was essential to ensure that study data were of sufficiently high quality for record linkage.

From a clinical perspective, the information provided about the family as a whole is more important than the accuracy of individual reports. The observation in this study that only two of six families who actually met surveillance criteria were identified at interview is a particular concern, implying that reliance on interview data in a clinical context could result in many families who actually meet criteria for significant family history being overlooked. Conversely, of five families reported at interview to meet the chosen criteria, only two were confirmed by record linkage to meet this classification. In practice, such an effect could lead to surveillance being

applied unnecessarily. While the numbers involved are too low to provide a conclusive assessment of the clinical utility of family history information reported at interview, our results illustrate that incomplete or inaccurate interviewee reporting could have a substantial impact on genetic risk assessment.

The appropriate family history criteria for offering genetic counselling, colonoscopic surveillance, or genetic testing is the subject of much current debate, and is likely to remain so. The findings of our study are highly relevant to this discussion, as they suggest that family history information obtained by interview may be misleading, and that verification of both positive and negative interviewee reports should be conducted whenever possible.

ACKNOWLEDGEMENTS

We acknowledge with thanks the considerable expertise of Norma Brown in interviewing the subjects in this study and the follow up genealogy by Rhona de Mey and Alison Fordyce of MRC Registry. We also thank Susan Frame and James Boyd for carrying out the record linkage in ISD, and Niall Anderson for guidance regarding statistical methods. RJM holds a Scottish Health Department Chief Scientist's Office PhD studentship. The work was supported by Cancer Research Campaign (now Cancer Research UK) Program Grant to AHW and CCB-SP1370/0501.

Authors' affiliations

R J Mitchell, H Campbell, Public Health Sciences, Department of Community Health Sciences, University of Edinburgh, Edinburgh, UK
D Brewster, Information and Statistics Division (ISD), Common Services Agency for NHS Scotland, Trinity Park House, Edinburgh, UK
M E M Porteous, Department of Clinical Genetics, University of Edinburgh, Western General Hospital, Edinburgh, UK
A H Wyllie, C C Bird, Department of Pathology, University of Edinburgh, Edinburgh, UK
M G Dunlop, Colon Cancer Genetics Group, University of Edinburgh, Division of Clinical and Molecular Medicine and MRC Human Genetics Unit, Western General Hospital, Edinburgh, UK

REFERENCES

- 1 **Johns LE, Houlston RS.** A systematic review and meta-analysis of familial colorectal cancer risk. *Am J Gastroenterol* 2001;**96**:2992-3003.
- 2 **Dunlop MG.** British Society for Gastroenterology, The Association of Coloproctology for Great Britain and Ireland. Guidance on gastrointestinal surveillance for hereditary non-polyposis colorectal cancer, familial

- adenomatous polyposis, juvenile polyposis, and Peutz-Jeghers syndrome. *Gut* 2002;**51**(suppl 5):V21-7.
- 3 **Dunlop MG.** British Society for Gastroenterology, The Association of Coloproctology for Great Britain and Ireland. Guidance on large bowel surveillance for people with two first degree relatives with colorectal cancer or one first degree relative diagnosed with colorectal cancer under 45 years. *Gut* 2002;**51**(suppl 5):V17-20.
- 4 **Glanz K, Grove J, Le Marchand L, et al.** Underreporting of family history of colon cancer: correlates and implications. *Cancer Epidemiol Biomarkers Prev* 1999;**8**:635-9.
- 5 **Kerber RA, Slattery ML.** Comparison of self-reported and database-linked family history of cancer data in a case-control study. *Am J Epidemiol* 1997;**146**:244-8.
- 6 **Floderus B, Barlow L, Mack TM.** Recall bias in subjective reports of familial cancer. *Epidemiology* 1990;**1**:318-21.
- 7 **Love RR, Evans AM, Josten DM.** The accuracy of patient reports of a family history of cancer. *J Chronic Dis* 1985;**38**:289-93.
- 8 **Sijmons RH, Boonstra AE, Reefhuis J, et al.** Accuracy of family history of cancer: clinical genetic implications. *Eur J Hum Genet* 2000;**8**:181-6.
- 9 **Douglas FS, O'Dair LC, Robinson M, et al.** The accuracy of diagnoses as reported in families with cancer: a retrospective study. *J Med Genet* 1999;**36**:309-12.
- 10 **Koch M, Hill GB.** Problems in establishing accurate family history in patients with ovarian cancer of epithelial origin. *Cancer Detect Prev* 1987;**10**:279-83.
- 11 **Katballe N, Juul S, Christensen M, et al.** Patient accuracy of reporting on hereditary non-polyposis colorectal cancer-related malignancy in family members. *Br J Surg* 2001;**88**:1228-33.
- 12 **Aitken J, Bain C, Ward M, et al.** How accurate is self-reported family history of colorectal cancer? *Am J Epidemiol* 1995;**141**:863-71.
- 13 **Novakovic B, Goldstein AM, Tucker MA.** Validation of family history of cancer in deceased family members. *J Natl Cancer Inst* 1996;**88**:1492-3.
- 14 **Kee F, Collins BJ.** How prevalent is cancer family syndrome? *Gut* 1991;**32**:509-12.
- 15 **Bondy ML, Strom SS, Colopy MW, et al.** Accuracy of family history of cancer obtained through interviews with relatives of patients with childhood sarcoma. *J Clin Epidemiol* 1994;**47**:89-96.
- 16 **Kuijten RR, Bunin GR, Meadows T.** Validation and accuracy of cancer reporting in first-degree relatives (SER abstract). *Am J Epidemiol* 1992;**136**:1019.
- 17 **Newcombe HB.** *Handbook of record linkage.* New York: OUP, 1988.
- 18 **Heasman MA, Clarke JA.** Medical record linkage in Scotland. *Health Bull (Edinb)* 1979;**37**:97-103.
- 19 **Kendrick SW, Clarke JA.** The Scottish Medical Record Linkage System. *Health Bull (Edinb)* 1979;**51**:72-9.
- 20 **ISD (Information and Statistics Division of the Scottish Executive).** The Scottish Record Linkage System (ISD information document). Edinburgh: NHS Scotland, 2002.
- 21 **Wilson EB.** Probable inference, the law of succession, and statistical inference. *JASA* 1927;**22**:209-12.
- 22 **Agresti A, Coull BA.** Approximate is better than "exact" for interval estimation of binomial proportions. *Am Stat* 1998;**52**:119-26.
- 23 **Brewster DH, Crichton J, Harvey JC, et al.** Completeness of case ascertainment in a Scottish regional cancer registry for the year 1992. *Public Health* 1997;**111**:339-43.
- 24 **Brewster D, Muir C, Crichton J.** Registration of colorectal cancer in Scotland: an assessment of data accuracy based on review of medical records. *Public Health* 1995;**109**:285-92.