The reliability of cancer registry records

In their recent paper, which compared information obtained from the cancer registry with data abstracted from medical records, Pollock and Vickers concluded that “disagreements over date of diagnosis will bias survival data, disagreements over site will affect incidence data and trends, and disagreements over treatment undermine the case for using registry data to evaluate care.” To what extent are these conclusions supported by data included in their paper?

For cases with data available both from the cancer registry and from case notes, dates of diagnosis agreed to within 30 days or less for 305 of 385 cases (79%). For date of death, there was exact agreement between the two data sets for 245 of 270 cases (91%). These findings show some lack of precision in recording date of diagnosis, particularly in cases for which registration was initiated from the death certificate. This paper has not presented data which show systematic bias from misclassification of date of diagnosis or date of death, either for the sample as a whole or for subgroups. In our own study of bladder cancer we found that there was exact agreement between data in the cancer registry and case notes for date of first operation in 83% of cases. For discordant cases the median difference was –1 day (interquartile range –10 to 3 days). Similarly for date of death there was exact agreement in 93% of cases and for discordant cases the median difference was 1–3 days. These data showed a high level of agreement between case notes and cancer registry data. In a small proportion of cases the date of first operation or date of death were recorded imprecisely but there was no evidence of a systematic bias that might influence survival estimates.

Turning to data presented for tumour site, the rate of discordance is similar to that reported for other registries. After excluding cases which the reviewers were unable to classify topographically (which are difficult to interpret as the authors do not give the codes used to sample from the cancer registry), data given in their table 4 are consistent with a kappa statistic of 0.86 (95% confidence interval 0.70 to 0.96). According to criteria conventionally used to evaluate the quality of epidemiological data, this is consistent with excellent agreement between the two data sources. No data were presented to show that there is any significant recording of a death changed over time in a manner that might influence the assessment of secular trends; in fact, the rate of discordance was the same in both years studied.

It is perhaps clear that cancer registries are not currently in a position to document cancer treatment in a comprehensive way, this does not undermine the case for using cancer registrations as a sampling frame for evaluative research. Cancer registry data have been shown repeatedly to be of value in comparing the outcome of cancer at international, national, and local levels. Recognition of a central role for population-based cancer registries in public health research has led to the increasing use of registers in the study of other conditions.

Several other points require qualification. The authors seem to regard their own data as confirming material to the effect that they have not evaluated the reliability of their own data abstraction and it is possible that an experienced cancer registry clerk may perform better than a less experienced researcher. Because the authors only retrieved case notes for 62% of cases provided by the cancer registry, reliability was studied for 416 cases and not the 673 cases mentioned in the title of the paper. The term “death certificate only registration” is used inconsistently and the term is used in the abstract to include cases for which details of diagnosis and treatment were available from the registry.

Cancer registrations, in common with other sources of routinely collected information, are not free from error and every effort should be made to allow cancer registries to improve the quality of their work. To observe that errors exist is unremarkable, it is more important to estimate the impact of such errors on subsequent data analyses. Evaluation of cancer registry data should be performed with care. The conclusion of this paper receive only limited support from the data presented.

MARTIN GULLIFORD
Department of Public Health Medicine, UMDS, Guy’s and St Thomas’ Maugle and Dental School, St Thomas’ Hospital, London SE1 7EH