How Do Cancer Registries in Europe Estimate Completeness of Registration?

I. Schmidtmann; M. Blettner
Institute of Medical Biostatistics, Epidemiology and Informatics, University of Mainz, Mainz, Germany

Keywords
Cancer registry, quality indicators, survey

Summary
Objectives: Several methods for estimating completeness in cancer registries have been proposed. Little is known about their relative merits. Before embarking on a systematic comparison of methods we wanted to know which indicators were currently in use and whether there had been comparative investigations of estimation methods.
Methods: We performed a survey among European cancer registries asking which methods for estimating completeness they used and whether they had performed comparisons of methods.
Results: One hundred and ninety-five European cancer registries were contacted after identification using membership directories of the European Network of Cancer Registries (ENCR) and of the International Association of Cancer Registries (IACR). Fifty-six (29%; 22%–36%) of the 195 cancer registries replied. Forty-eight (86%; 74%–94%) of these stated that they estimated completeness. Thirty-five (73%; 58%–85%) used historic comparisons, 31 (65%; 49%–78%) compared their data with a reference registry, 28 (58%; 43%–72%) registries used mortality incidence ratio. Capture-recapture methods were applied in only 12 (25%; 14%-40%) registries. The flow method was used by ten (21%; 10%–35%) registries. There were regional differences in the use of methods.
Comparisons of methods were rare and usually restricted to real data at hand. A systematic comparison including all indicators actually in use in cancer registries was not reported.
Conclusions: A comparison of methods under well defined realistic conditions seems to be indicated. Unifying the methods for estimating completeness would improve validity of comparisons between cancer registries.

Correspondence to:
Irene Schmidtmann
Institute of Medical Biostatistics, Epidemiology and Informatics
University of Mainz
55101 Mainz
Germany
E-mail: schmidtmann@imbei.uni-mainz.de

Methods Inf Med 2009; 48: 267–271
doi: 10.3414/ME0559
received: April 4, 2008
accepted: September 20, 2008
prepublished: March 31, 2009

Introduction

The purpose of population-based cancer registries is to estimate the cancer burden in their area, to observe trends and regional differences and to provide a database for epidemiological research. Population-based cancer registries can only reach their goals when they are complete or to a large extent complete. An incomplete cancer registry is only of limited use. First, the cancer incidence is underestimated. Second, it is also possible that some subgroups of cancer patients are more likely to be registered than others, thereby misrepresenting the distribution of age or stage of disease or the regional distribution of cases. Third, trends in completeness of registration may be mistaken for trends in incidence. Researchers can assess whether cancer registry data are representative and useful for their purpose if an estimate of completeness of registration is available. Several of the German cancer registries still have to show that they operate effectively. One of the indicators used to measure effectiveness is completeness of registration. Outside Germany and the United Kingdom not many cancer registries publish indicators for completeness regularly.

Available Methods for Estimating Completeness

Completeness is defined as the proportion of diagnosed cancer cases that are registered. Several methods to assess completeness have been proposed [1].

1. Completeness can be estimated using the proportion of cases for which the first and possibly only source of information is a death certificate (DCN cases). An estimate of completeness is then given by

$$I_1 = \frac{1}{(1-\mathrm{DCN}) + \left(\frac{\mathrm{DCN}}{\mathrm{M:I}}\right)}$$

   where DCN denotes the proportion of DCN cases and M:I denotes the mortality to incidence ratio [1] which can be computed as the number of deaths from a particular cancer in a given year divided by the number of incident cases for the same cancer in this year.
2. Comparing current incidence rates or numbers of cases registered to appropriate numbers from the past within the same registry is termed historic data method [1]. Observed trends can be taken into account.
3. Completeness can also be estimated by comparing figures from the cancer registry under consideration with data from a presumably complete reference registry. This can be done by simply comparing incidence rates.
   a) In a more rigorous approach, the number of cases can be estimated using

age-specific incidence rates of the reference registry and the demographic structure of the registry population under consideration.

b) The expected number of cases is estimated from the mortality rates in the registry population under consideration and the incidence and mortality rates in the reference registry population.

4. Capture-recapture methods have been used to estimate completeness of disease registries [2–5], although their usefulness has been questioned [6, 7] and some problems are known [8].
5. Independent case ascertainment [1] can be regarded as a special case of capture-recapture methods. An independent data base of cancer cases may be available, e.g. data collected for a clinical or epidemiological study or data obtained for administrative purposes. This database and the registry database can be linked – given confidentiality issues do not prohibit the linkage. The proportion of cancer cases in the independent database that are also known to the cancer registry can serve as an estimate of completeness.
6. The flow method by Bullard et al. [9] uses three time-dependent probabilities: the probability that a patient survives until time t, the probability that the death certificate of a patient who dies in the interval after time t mentions cancer, and the probability that a patient survives until time t and remains unregistered. From these the completeness at any time can be estimated.
7. The method proposed by Haberland [10, 11] is based on a paper by Colonna et al. [12] in which cancer incidence rates for the whole of France were estimated from incidence rates from regional registries covering only part of the country and both national and regional mortality rates. This method is also based on mortality incidence ratios (M/I) but extends the simple rule of proportion by introducing log-linear models for trends of mortality and incidence.

The proposed methods have as prerequisites certain assumptions such as independence of sources, constant trend, or similarity of cancer risk and prognosis in compared regions. However, these assumptions are rarely, if ever, met in practice. Little is known about the merits and weaknesses of these procedures in realistic settings. Mattauch [13] performed a comparison of methods using data from the Münster cancer registry and found considerable differences between methods in some instances. Silcocks and Robinson [14] compared the flow method and capture-recapture methods in a simulation study with focus on the confidence intervals obtained. They found that point estimates for completeness differed although width of confidence intervals was similar. They concluded that one or possibly both approaches may not be appropriate and advocated the use of the flow method. Recently they also compared the flow method and the DCN method using Ajiki's formula [15, 16] in a simulation study [17]. They found that the flow method gave more realistic results than the DCN method which grossly underestimated completeness.

We wanted to obtain an overview over methods currently in use in cancer registries in Europe. The main objective of the research reported here was to find out whether cancer registries a) actually estimated completeness, b) if so which methods they used and how frequently, and c) if not why not. Additionally, we were interested in the software used and in modifications of published methods. A further aim was to collect information on unpublished methodological and comparative investigations. The survey was intended as a starting point for a more methodological research project. An in-depth discussion of merits and weaknesses of the proposed estimation methods is beyond the scope of this paper. The latter is addressed in [18].

Table 1 Response rates by region

Region of registry	Responders		Number of registries in region
	N	%	
East (Belarus, Bulgaria, Czech Republic, Hungary, Poland, Republic of Moldova, Romania, Russian Federation, Slovakia, Ukraine; Armenia, Georgia, Kyrgyzstan, Ukraine)	5	18	28
North (Channel Islands, Denmark, Estonia, Faeroe Islands, Finland, Iceland, Ireland, Isle of Man, Latvia, Lithuania, Norway, Sweden, United Kingdom of Great Britain and Northern Ireland; Bermuda)	15	50	30
South (Albania, Andorra, Bosnia and Herzegovina, Croatia, Gibraltar, Greece, Holy See, Italy, Malta, Montenegro, Portugal, San Marino, Serbia, Slovenia, Spain, The former Yugoslav Republic of Macedonia; Cyprus, Turkey)	9	13	69
West (Austria, Belgium, France, Germany, Liechtenstein, Luxembourg, Monaco, Netherlands, Switzerland)	27	40	68
All regions	56	29	195

Methods

For practical purposes the survey was restricted to European cancer registries, European being defined by membership in the ENCR or by being mentioned in the European section of the IACR [19].

A short questionnaire was developed, mainly including questions on methods used for reporting and calculating completeness. We asked about the DCN-method, historical comparison, comparison with a reference registry, independent case ascertainment, flow method [9], method based on mortality incidence ratio as described by Colonna et al. [12], and the capture-recapture method. Further methods could be specified. We also asked for the frequency of use (never, routinely, on special occasion). As several respondents ticked the boxes "routine use" or "on

special occasion" but left the others empty, we interpreted missing values as "never". Additionally we asked for availability of software, performance of method comparisons, references, contact details and interest in feedback.

We first distributed the questionnaire among the 11 German population-based cancer registries. After receiving responses from German registries we changed the layout slightly. In April 2005, we sent questionnaires to all 153 non-German cancer registries listed in [20] and asked for completion. The questionnaires were sent with a cover letter that explained the research project using stationery of the Cancer Registry Rhineland Palatinate. In August 2006, the list was updated using the membership directories of the ENCR [21] and the European section of the IACR [19]. We contacted all registries in the updated list containing 195 addresses, unless they had already returned the questionnaire. This reminder was sent by E-mail when an E-mail address was available, otherwise by ordinary mail.

The returned questionnaires were entered into an Access database. The analysis is based on all questionnaires received up to October 2006. We computed appropriate absolute and relative frequencies. We also give 95% exact confidence intervals for relative frequencies. Statistical analysis was performed using SAS 9.1 after converting the Access database into a SAS data set.

For regional analyses, the group definition by the United Nations Population Division [22] was adapted as some of the ENCR members belong to Asia or America according to this classification. Asian countries formerly belonging to the Soviet Union are added to the Eastern Europe category, Turkey and Cyprus are added to Southern Europe category, and Bermuda is combined with the United Kingdom (▶see Table 1).

Results

Only 56 registries (29%; 22%–36%) returned the questionnaire. The response rate was higher in Northern and Western Europe.

Forty-eight (86%; 74%–94%) of these registries stated that they estimated completeness, 8 (14%; 6%–26%) reported that they did not estimate completeness. Three of the latter deemed it unnecessary. The following other reasons for not estimating completeness were stated by one or two registries each: takes too much time, no software available, nobody in the registry capable of doing it, incidence comparison with other registries is performed, all pathologists contribute, other priorities and limited staffing, new registry. The proportion of registries using the various methods can be seen in ▶Table 2.

Table 2 Frequency of use of various methods (multiple answers possible)

Method	Method is applied						Number of registries
	no information or never		routine use		special occasion		
	N	%	N	%	N	%	N
Bullard	38	79	5	10	5	10	48
Capture/recapture	36	75	5	10	7	15	48
DCN method	22	46	14	29	2	4	48
Historical comparison	15	31	35	73	0	0	48
independent case ascertainment	28	58	9	19	13	27	48
Mortality/incidence ratio (Colonna)	15	31	18	38	10	21	48
Comparison with reference registry	19	40	24	50	7	15	48

Most registries reported that they used historical comparisons routinely. Many registries compared their data with a reference registry, mainly as a routine exercise. Mortality incidence ratio was a commonly used indicator, whereas independent case ascertainment and other capture-recapture methods tended to be used on special occasions. The flow method was only used in ten registries.

There were regional differences in the choice of method (▶see Fig. 1). While most responding cancer registries in Southern and Northern Europe (n = 18, 86%; 64%–97%) reported the use of historical comparisons only 13 (59%; 36%–79%) of the cancer registries in Western Europe did so. The DCN-method seems to be popular in Eastern Europe – four of five replying registries used it – whereas in Western Europe only four (18%; 5%–40%) of the responding registries mentioned using it. The flow method and capture-recapture methods were not used in the responding Eastern European registries at all, whereas the flow method was used by six (43%; 18%–71%) of the Northern European registries – particularly in the United Kingdom.

Few registries confirmed the availability of software. Twelve (25%; 14%–40%) cancer registries stated to have software for comparison with a reference registry, ten (21%; 10%–35%) of the registries quoted software for historical comparisons or for mortality incidence ratio. Eight (17%; 7%–30%) registries had access to software for the flow method. Software for the DCN method was mentioned by six (13%; 5%–25%) registries, independent case ascertainment and capture-recapture by four (8%; 2%–20%) and other methods by three (6%; 1%–17%).

Modifications of published methods were rarely mentioned. Two registries compared current M/I ratios with M/I ratios from previous years. One registry explicitly stated to compare M/I ratio with those from other registries.

Estimation of completeness was mostly performed by epidemiologists (n = 30, 63%; 47%–76%) or by statisticians (n = 23, 48%; 33%–63%). Medical doctors (n = 13, 27%; 15%–42%), computer scientists (n = 10, 21%; 10%–35%) and cancer registrars (n = 7, 15%; 6%–28%) were also involved.

Only 14 (29%; 17%–44%) of the cancer registries that performed estimation of completeness had ever applied more than one method to a dataset and compared results.

When asked for publications 21 (44%; 29%–59%) cancer registries referred to tech-

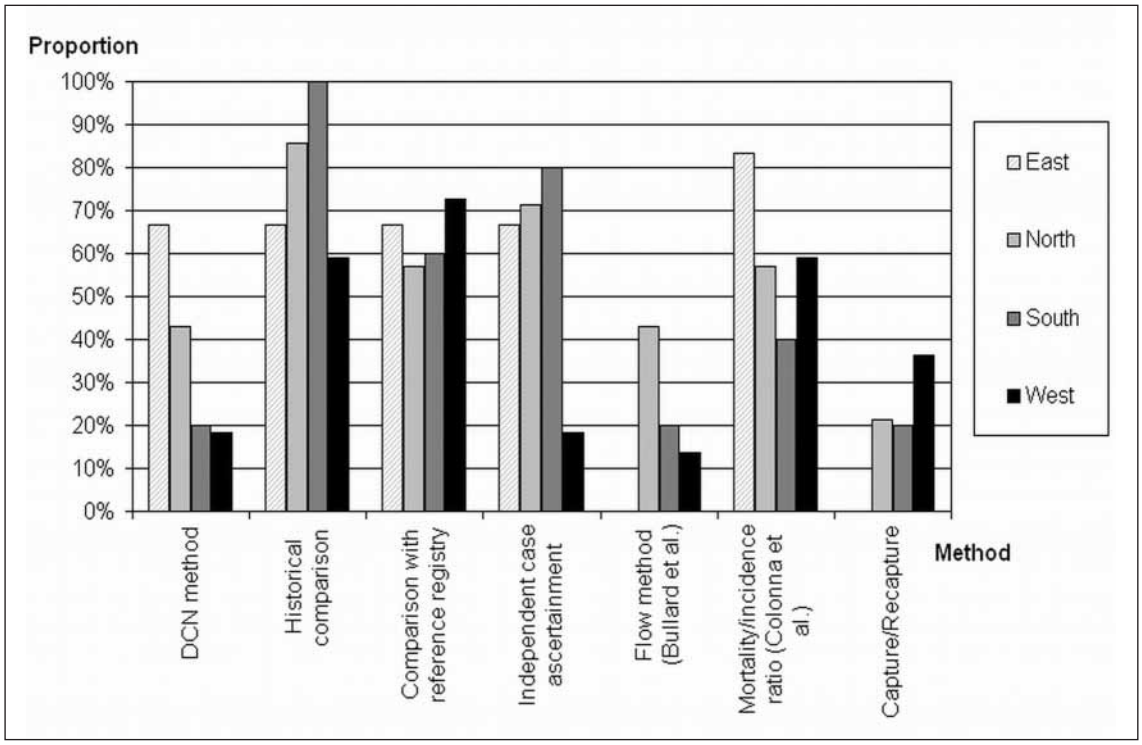

Fig. 1 Proportion of cancer registries within each region using various methods of estimating completeness

nical reports, 11 (23%; 12%–37%) mentioned publication in peer-reviewed journals and five (10%; 3%–23%) other publications. Fifteen (31%; 19%–46%) registries answered that they had not published results concerning estimation of completeness.

Some registries provided useful references to their own work concerning estimation of completeness or gave information on ongoing work in this field, e.g. [14, 23–30].

Discussion

Summary of Results

Most of the cancer registries that completed the questionnaire estimate completeness in some way. Simple methods such as DCN-method, historical comparisons, comparison with reference registry and M/I ratio are preferred. Modification or extension of published methods are rarely stated. Within the cancer registries methodological developments and comparisons are rare and if they have been performed they have been based on real registry data at hand.

Limitations

The generalization of our results is limited as the response rate is low and varies between European regions. A fairly large proportion of cancer registries from Germany, the Netherlands, Scandinavia and the United Kingdom responded. These cancer registries are however generally reasonably well funded and cover major parts of the population. On the other hand, response rates are very low for registries in Eastern and Southern Europe, many of which are small, restricted to a small population or to a small selection of diseases, or have to function on limited funds. The more sophisticated methods tend to be used more frequently in Western Europe and Northern Europe. It may be speculated that small registries and registries with little resources have not responded. Such registries presumably would prefer simple methods to estimate completeness or not estimate it at all. This would mean that the use of the more sophisticated procedures was overestimated.

Although there had been a pre-test, sometimes questions where not understood as they had been intended. Some registries reported plans rather than the current status. From further comments given, we concluded that sometimes by "method based on M/I ratio" only the simple M/I ratio instead of the model-based procedure was meant. The availability of software was often negated although the respective method was used. From free text notes it may be concluded that in these cases standard software such as Excel or statistical packages such as STATA or SAS are used.

Conclusion

This survey confirms the impression gained from literature search: there are very different methods in use and there are hardly any comparative studies about the performance of indicators of completeness. The investigations found in the literature and indicated in this survey apply several indicators for completeness to the same real dataset. However, it is unsatisfactory that applying several indicators may yield considerably different results [13]. A systematic comparison including all indicators actually in use in cancer registries has not been performed so far. So a method comparison under realistic and well defined conditions extending or complementing the work by Silcocks and Robinson [14, 17] seems to be indicated. Unifying the methods for estimating completeness would improve validity of comparisons of completeness between

cancer registries. This can be targeted once the relative merits of the procedures in use have been assessed.

Acknowledgments

The authors wish to thank all colleagues who took the time to complete the questionnaire. The technical assistance of Ulrike Knoll, Dagmar Lautz, and Lamia Yousif is gratefully acknowledged. We would also like to thank Claudia Spix for critically reading and commenting on an earlier draft of this paper.
This work contains part of the first author's PhD thesis under preparation.

References

1. Parkin DM, Chen VW, Ferlay J, Galceran J, Storm HH, Whelan SL. Comparability and Quality Control in Cancer Registration. Lyon: IARC Press; 1994.
2. Robles SC, Marrett LD, Clarke EA, Risch HA. An application of capture-recapture methods to the estimation of completeness of cancer registration. J Clin Epidemiol 1988; 41 (5): 495–501.
3. Brenner H, Stegmaier C, Ziegler H. Estimating completeness of cancer registration in Saarland/Germany with capture-recapture methods. Eur J Cancer 1994; 30A (11): 1659–1663.
4. Brenner H, Stegmaier C, Ziegler H. Estimating completeness of cancer registration: an empirical evaluation of the two source capture-recapture approach in Germany. J Epidemiol Community Health 1995; 49 (4): 426–430.
5. Berghold A, Stronegger WJ, Wernecke KD. A model and application for estimating completeness of registration. Methods Inf Med 2001; 40 (2): 122–126.
6. Schouten LJ, Straatman H, Kiemeney LA, Gimbrere CH, Verbeek AL. The capture-recapture method for estimation of cancer registry completeness: a useful tool? Int J Epidemiol 1994; 23 (6): 1111–1116.
7. Tilling K. Capture-recapture methods – useful or misleading? Int J Epidemiol 2001; 30 (1): 12–14.
8. Brenner H. Application of capture-recapture methods for disease monitoring: potential effects of imperfect record linkage. Methods Inf Med 1994; 33 (5): 502–506.
9. Bullard J, Coleman MP, Robinson D, Lutz JM, Bell J, Peto J. Completeness of cancer registration: a new method for routine use. Br J Cancer 2000; 82 (5): 1111–1116.
10. Haberland J, Bertz J, Görsch B, Schön D. Krebsinzidenzschätzungen für Deutschland mittels log-linearer Modelle. Gesundheitswesen 2001; 63 (8–9): 556–560.
11. Haberland J, Schön D, Bertz J, Görsch B. Vollzähligkeitsschätzungen von Krebsregisterdaten in Deutschland. Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz 2003; 46 (9): 770–774.
12. Colonna M, Grosclaude P, Faivre J, Revzani A, Arveux P, Chaplain G, et al. Cancer registry data based estimation of regional cancer incidence: application to breast and colorectal cancer in French administrative regions. J Epidemiol Community Health 1999; 53 (9): 558–564.
13. Mattauch V. Strukturanalyse des Epidemiologischen Krebsregisters für den Regierungsbezirk Münster: Identifikation von Schwachstellen und Konzepte für eine kontinuierliche Qualitätssicherung. Münster: Inauguraldissertation. Medizinische Fakultät der Westfälischen Wilhelms-Universität Münster; 2005.
14. Silcocks PB, Robinson D. Completeness of ascertainment by cancer registries: putting bounds on the number of missing cases. J Public Health 2004; 26 (2): 161–167.
15. Ajiki W, Tsukuma H, Oshima A. Index for evaluating completeness of registration in population-based cancer registries and estimation of registration rate at the Osaka Cancer Registry between 1966 and 1992 using this index (in Japanese). Nippon Koshu Eisei Zasshi 1998; 45 (10): 1011–1017.
16. Kamo K, Kaneko S, Satoh K, Yanagihara H, Mizuno S, Sobue T. A mathematical estimation of true cancer incidence using data from population-based cancer registries. Jpn J Clin Oncol 2007; 37 (2): 150–155.
17. Silcocks PB, Robinson D. Simulation modelling to validate the flow method for estimating completeness of case ascertainment by cancer registries. J Public Health 2007; 29 (4): 455–462.
18. Schmidtmann I. Estimating completeness in cancer registries – comparing capture-recapture methods in a simulation study. Biom J 2008; 50 (6): 1077–1092.
19. International Association of Cancer Registries. IACR Membership List. International Association of Cancer Registries Homepage 2006 (cited 2006 Aug 9). Available from: URL: http://www.iacr.com.fr
20. Tyczynski JE, Demaret E, Parkin DM. Standards and Guidelines for cancer registration. Lyon: IARC Press; 2003.
21. European Network of Cancer Registries. ENCR Membership List. European Network of Cancer Registries Homepage 2006 (cited 2006 Aug 9). Available from: URL: http://www.encr.com.fr/
22. UN Population Division. Definition of Major Areas and Regions. UN Population Division Homepage 2007 (cited 2007 Aug 10). Available from: URL: http://esa.un.org/unpp/index.asp?panel=5
23. Crocetti E, Miccinesi G, Paci E, Zappa M. An application of the two-source capture-recapture method to estimate the completeness of the Tuscany Cancer Registry, Italy 4. Eur J Cancer Prev 2001; 10 (5): 417–423.
24. Schouten LJ, Hoppener P, van den Brandt PA, Knottnerus JA, Jager JJ. Completeness of cancer registration in Limburg, The Netherlands. Int J Epidemiol 1993; 22 (3): 369–376.
25. Brewster DH, Crichton J, Harvey JC, Dawson G. Completeness of case ascertainment in a Scottish regional cancer registry for the year 1992. Public Health 1997; 111 (5): 339–343.
26. Lund E. Pilot study for the evaluation of completeness of reporting to the cancer registry. Incidence of Cancer in Norway 1978. Oslo: The Cancer Registry of Norway; 1981. pp 11–14.
27. Helseth A, Langmark F, Mork SJ. Neoplasms of the central nervous system in Norway. I. Quality control of the registration in the Norwegian Cancer Registry. APMIS 1988; 96 (11): 1002–1008.
28. Mork J, Thoresen S, Faye-Lund H, Langmark F, Glattre E. Head and neck cancer in Norway. A study of the quality of the Cancer Registry of Norway's data on head and neck cancer for the period 1953-1991. APMIS 1995; 103 (5): 375–382.
29. Harvei S, Tretli S, Langmark F. Quality of prostate cancer data in the cancer registry of Norway. Eur J Cancer 1996; 32A (1): 104–110.
30. Tingulstad S, Halvorsen T, Norstein J, Hagen B, Skjeldestad FE. Completeness and accuracy of registration of ovarian cancer in the cancer registry of Norway1. Int J Cancer 2002; 98 (6): 907–911.