available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Interpreting international comparisons of cancer survival: The effects of incomplete registration and the presence of death certificate only cases on survival estimates

David Robinson[a,*], *Risto Sankila*[b], *Timo Hakulinen*[b], *Henrik Møller*[a]

[a]*King's College London, Thames Cancer Registry, 1st Floor, Capital House, 42 Weston Street, London SE1 3QD, UK*
[b]*Finnish Cancer Registry, Institute for Statistical and Epidemiological Cancer Research, Liisankatu 21 B, FIN-00170 Helsinki, Finland*

ARTICLE INFO

Article history:
Received 19 December 2006
Accepted 2 January 2007
Available online 14 February 2007

Keywords:
Cancer registries
Relative survival
Death certificate only
Incompleteness

ABSTRACT

We have assessed the impact on survival estimates based on cancer registry data of incomplete ascertainment of cancer cases and the presence of cases registered purely from death certificate information (DCO cases). Using data from the Thames and Finnish Cancer Registries we obtained five-year relative survival estimates for 12 cancer sites, excluding DCOs as usual. We then made adjustments to allow for the effects of both the known proportion of DCOs and the estimated proportion of missing cases for each site. In general, adjusting for DCOs led to lower survival estimates, whilst adjusting for incompleteness had the opposite effect. The Finnish data were largely complete and had small proportions of DCOs, and hence the adjustments had little effect on estimated survival. The changes in the Thames estimates were more marked. When performing cohort survival analysis (based on diagnoses between 1990 and 1994), the increases in the survival estimates gained from adjusting for incompleteness were for the most part offset by the decrease produced when adjusting for DCOs. However, when performing period survival analysis based on the period 1997–2001 (when the DCO rate at Thames had fallen by around a half relative to the earlier period), the final estimates (adjusted for both effects) were generally higher than the unadjusted values – thus reducing the apparent difference between the two countries. It is important to take variations in DCO proportion and/or completeness into consideration when comparing survival estimates between different populations.

© 2007 Elsevier Ltd. All rights reserved.

1. Introduction

The EUROCARE studies[1–3] have reported substantial variation in cancer survival estimates between different European countries. For example, the most recent[3] showed the five-year age-standardised breast cancer relative survival for England to be some 8 percentage points lower than in Finland. For prostate cancer, the difference was almost 13 percentage points. However, it is unclear to what extent the observed differences are influenced by artefacts related to the acquisition of data by the cancer registries within the different countries.

Two factors affecting survival estimates based on cancer registry data are incompleteness of case ascertainment and the presence of death certificate only (DCO) cases. These two factors are related in a complex way. DCOs are cases where the registration is made purely from a death certificate in which cancer is mentioned, but for which no supporting clinical information is available and for whom no definitive date of diagnosis is known. Thus, a high proportion of DCOs

* *Corresponding author:* Tel.: +44 207 378 7688; fax: +44 207 378 9510.
E-mail address: dave.robinson@kcl.ac.uk (D. Robinson).
0959-8049/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2007.01.007

indicates incomplete ascertainment of cases. Improving data capture from clinical data sources will both increase overall case ascertainment and reduce DCOs. Improving trace-back of DCOs will reduce the DCO percentage, but have no effect on the overall completeness of ascertainment. Hence, while a high proportion of DCOs always implies incomplete ascertainment, a low DCO percentage does not necessarily indicate complete ascertainment.[4]

Incompleteness and the presence of DCOs will tend to bias calculated estimates in opposite directions.[5] For example, registries in the UK have routine access to death certificates, and hence their databases are largely complete with respect to fatal cancer cases. Where ascertainment is incomplete, the missing cases tend to be the long-term survivors, and this then biases estimates in the direction of poor survival. On the other hand, DCO cases have in general worse than average survival. This was demonstrated by Pollock and Vickers,[6] who found that five-year survival for colorectal cancer decreased by four to nine percentage points across four districts in the Thames Cancer Registry area in the UK after including DCOs for whom a date of diagnosis had been ascertained from case notes. In a subsequent study[7] they showed that DCO registration was associated with increasing age, decreasing survival, district of residence and place of death. Hence, the routine exclusion of such cases will bias survival estimates upwards.

Brenner and Hakulinen[8] have assessed the impact of underascertainment on relative survival rates under a number of theoretical scenarios. They showed that the magnitude of potential bias is dependent on both the time period affected by underascertainment and the type of survival analysis undertaken.

In this study, we have attempted to quantify the effects of DCOs and incomplete ascertainment more directly, by comparing results from the Finnish and Thames Cancer Registries, both before and after making appropriate adjustments (for incompleteness, DCO registrations and both).

2. Materials and methods

Cohort relative survival analyses were performed on datasets of cases from the two registries diagnosed between 1990 and 1994, with survival calculated up to the end of 1999. These analyses correspond to those published in EUROCARE-3. Sites studied were head and neck (ICD-10 codes C00-C14/C30-C32/C73), oesophagus (C15), stomach (C16), colon (C18), rectum (C19–C21), lung (C33–C34), skin melanoma (C43), breast (C50), cervix uteri (C53), corpus uteri (C54), ovary (C56) and prostate (C61).

Initial ('unadjusted') 5-year relative survival estimates were calculated, excluding DCO cases as usual. The datasets were then adjusted by assuming that the survival of the DCO cases is the same as the median survival of those cases who were originally registered from death certificates but subsequently successfully traced (matched for site, sex and age at death), and 'DCO adjusted' survival estimates calculated.

For each combination of registry, sex and cancer site, the Thames completeness program[9] (a model which simulates the process of ascertainment of cancer cases) was used to provide estimates of completeness of registration at five years after diagnosis, based on cases diagnosed in 1992 and deaths occurring in 1996. These estimates were then used to enhance each sex/site group within each registry by the appropriate proportion, making the assumption that these 'missing' cases all survived for more than five years, and 'incompleteness adjusted' survival estimates were calculated.

Finally, survival estimates were calculated after adjustment for the effects of both DCOs and incompleteness.

The whole process was repeated using period relative survival analysis, based on the period 1997–2001. This method,[10] rather than following a specific cohort of patients, is based on follow-up during the specified period and produces more up-to-date estimates of survival.

All calculations were performed with the statistical package Stata,[11] using the procedures strel and strelperiod developed by Prof. Michel Coleman's group at the London School of Hygiene and Tropical Medicine.

3. Results

Tables 1 and 2 show, respectively, the total numbers of cases and the proportions of DCO cases at each registry for the sites studied, both for the period 1990–1994 (on which the cohort analysis was based) and for 1997–2001 (on which the period analysis was performed). DCO proportions were consistently low in Finland, ranging from under 1% for melanoma and female breast cancer to 6% for lung cancer. The proportions were much higher at Thames during the period 1990–1994 (ranging from 6 to 7% for melanoma to 32% for female stomach cancer), but had improved substantially – falling by about a half – by the later period.

Estimates of completeness of ascertainment at five years after diagnosis are shown in Table 3. These were assumed to remain constant throughout the period of the study. The Finnish data appear to be largely complete, with estimates ranging from around 96% completeness for prostate cancer to 100% for ovarian cancer. Data at Thames were less complete, with estimates ranging from around 78% (female melanoma) to 95% (female stomach cancer).

Tables 4 and 5 show the results of the cohort and period survival analyses, respectively, in relation to the different adjustments. For most sites, the unadjusted survival estimates were higher for Finland than for Thames – the exceptions being oesophageal cancer, lung cancer in men and cervical cancer.

For the unadjusted Finnish data, period survival estimates were greater than cohort estimates for all sites, with the largest increases occurring in cancers of the colon, prostate and ovary, and in female oesophageal and male rectal cancers. For Thames cases the changes were less pronounced, with the greatest increases occurring in prostate and rectal cancers, and a slight reduction in estimated five-year survival for lung cancer, male head and neck cancer and female oesophageal cancer.

Adjusting for DCOs had little effect on the Finnish estimates, but reduced the estimates at Thames by substantial

Table 1 – Numbers of diagnosed cases by cancer site and period

Site		Finland		Thames	
		1990–1994	1997–2001	1990–1994	1997–2001
Head and Neck	(F)	2248	2417	3134	3524
	(M)	2267	2579	4931	5616
Oesophagus	(F)	445	414	2777	2870
	(M)	589	631	3999	4528
Stomach	(F)	2433	1938	4470	3368
	(M)	2700	2270	7208	5770
Colon	(F)	3199	3594	11,869	11,514
	(M)	2282	2845	9871	10,760
Rectum	(F)	1837	2000	5538	5933
	(M)	1923	2215	6641	7502
Lung	(F)	2161	2583	15,555	15,253
	(M)	8655	7539	29,806	24,272
Melanoma	(F)	1257	1564	2885	3623
	(M)	1267	1569	1990	2803
Breast	(F)	13,980	18,044	41,787	46,458
Cervix uteri	(F)	717	804	3729	2912
Corpus uteri	(F)	2908	3438	5101	5746
Ovary	(F)	2789	2908	7267	7219
Prostate	(M)	8720	16,080	21,491	30,033

Table 2 – DCO proportions (%) by cancer site and period

Site		Finland		Thames	
		1990–1994	1997–2001	1990–1994	1997–2001
Head and Neck	(F)	1.5	1.1	10.7	6.7
	(M)	1.4	1.1	9.0	5.5
Oesophagus	(F)	4.0	1.9	22.6	12.3
	(M)	2.9	3.2	23.3	8.9
Stomach	(F)	3.4	2.7	32.4	17.2
	(M)	3.0	2.9	25.3	12.0
Colon	(F)	2.7	2.5	23.3	12.7
	(M)	2.8	2.5	20.4	10.2
Rectum	(F)	1.6	1.3	15.8	7.8
	(M)	1.4	1.1	13.3	5.2
Lung	(F)	6.0	5.4	30.1	15.1
	(M)	5.1	5.8	27.8	13.7
Melanoma	(F)	0.2	0.1	5.7	4.2
	(M)	0.3	0.2	7.3	4.1
Breast	(F)	0.3	0.3	11.4	5.7
Cervix uteri	(F)	1.3	0.8	9.1	5.7
Corpus uteri	(F)	0.8	0.8	11.0	3.7
Ovary	(F)	1.3	1.7	18.4	11.3
Prostate	(M)	1.7	1.0	17.3	8.2

amounts. On the other hand, adjusting for incompleteness resulted in substantial increases in the estimates at Thames (particularly for prostate, lung, oesophageal and stomach cancers), but only marginal changes in the Finnish estimates, with the largest effects seen for lung cancer and female oesophageal cancer.

4. Discussion

Survival estimates are generally considered to be an indication of the quality and effectiveness of cancer care, with low survival implying inadequate or sub-optimal treatment in the population concerned. However, there are many other

Table 3 – Estimated 5-year completeness (%) by cancer site

Site		Finland	Thames
Head and Neck	(F)	98.2	87.4
	(M)	97.2	84.4
Oesophagus	(F)	97.0	90.5
	(M)	99.6	89.4
Stomach	(F)	99.4	94.5
	(M)	99.8	91.2
Colon	(F)	99.7	86.6
	(M)	97.9	88.9
Rectum	(F)	99.6	92.8
	(M)	97.5	89.0
Lung	(F)	97.7	90.6
	(M)	98.6	93.3
Melanoma	(F)	99.1	77.5
	(M)	99.4	84.9
Breast	(F)	98.5	85.0
Cervix uteri	(F)	99.6	87.7
Corpus uteri	(F)	98.6	88.3
Ovary	(F)	100.0	91.7
Prostate	(M)	96.1	82.6

factors which can affect such estimates. Prior et al.[12] showed that published differences in survival could be due to the failure of cancer registries to ascertain all cases with advanced disease, and concluded that 'comparisons across populations which vary in accuracy and completeness of registration can be misleading and should be treated with some circumspection'.

In our comparison of data from Finland and the UK, we have shown that survival estimates are affected by both the presence of DCOs and underascertainment of cancer cases. After adjusting for these factors, we found that large changes to the estimates occurred for sites with a large proportion of DCOs and/or a large percentage missing. Some of the differences between the two countries remained after adjustment, but some disappeared or were substantially reduced, and hence are possibly not 'true' differences.

Ascertainment was higher in Finland. This is likely to be due in part to the fact that cancer is a statutorily notifiable disease in Finland, but not in the UK. The Finnish data were virtually complete, and had few DCO cases. Consequently, adjustment for these factors had little effect on survival estimates. In the Thames cohort analysis, the increases in the cohort survival estimates gained from adjusting for incompleteness were for the most part offset by the decrease produced when adjusting for DCOs. However, the decrease in the proportion of DCOs at Thames in the later period resulted in a much weaker bias due to this effect in the period analysis, where the final estimates (adjusted for both effects) were generally larger than the unadjusted estimates.

In most instances, the period survival estimates were higher than the corresponding cohort estimates, showing that survival has improved at these sites in both countries. This was confirmed by running a cohort analysis on Thames cases diagnosed in the later period (1997–2001). (Results not shown.)

The assumption that the 'missing' cases all survive for more than five years is a major one, and is less likely to be valid for the more fatal cancers (e.g. oesophagus, stomach and lung). These are the sites for which the effects of the adjustments are greatest, and the results should be treated with

Table 4 – Five-year relative survival estimates (%) by cancer site and adjustment factor – cohort analysis

Site		Finland				Thames			
		None	DCOs	Incomp	Both	None	DCOs	Incomp	Both
Head and neck	(F)	80.02	80.00	80.52	80.51	64.84	57.27	71.58	64.38
	(M)	64.89	64.87	66.27	66.25	57.68	52.07	67.66	62.20
Oesophagus	(F)	7.52	7.49	9.65	9.62	11.83	8.49	24.77	18.60
	(M)	6.40	6.39	7.03	7.02	7.92	5.67	22.62	16.96
Stomach	(F)	25.37	25.28	25.80	25.71	16.65	9.48	24.84	14.88
	(M)	22.46	22.39	22.61	22.54	15.49	10.46	28.01	19.92
Colon	(F)	50.09	49.97	50.35	50.23	47.10	33.76	59.02	44.94
	(M)	50.92	50.81	52.70	52.60	44.40	34.31	56.08	44.15
Rectum	(F)	51.42	51.36	51.61	51.55	48.54	39.83	54.53	45.54
	(M)	48.34	48.31	50.25	50.23	44.37	37.71	54.44	47.29
Lung	(F)	12.19	12.09	14.67	14.55	9.15	5.55	22.43	14.63
	(M)	8.36	8.31	10.13	10.07	9.08	5.79	19.40	12.97
Melanoma	(F)	86.24	86.27	86.50	86.54	86.25	81.35	91.00	87.06
	(M)	78.34	78.34	78.50	78.50	74.37	68.77	80.24	75.19
Breast	(F)	83.02	82.99	83.40	83.37	78.30	68.90	83.37	75.03
Cervix uteri	(F)	60.00	59.97	60.21	60.19	66.03	59.68	71.33	65.35
Corpus uteri	(F)	82.66	82.65	83.10	83.08	74.27	65.76	78.89	70.99
Ovary	(F)	47.22	47.16	47.22	47.16	36.26	28.47	43.81	35.31
Prostate	(M)	64.98	64.95	67.97	67.94	61.12	49.31	76.49	64.57

Table 5 – Five-year relative survival estimates (%) by cancer site and adjustment factor – period analysis

Site		Finland				Thames			
		None	DCOs	Incomp	Both	None	DCOs	Incomp	Both
Head and Neck	(F)	84.67	84.64	85.09	85.06	65.21	60.59	71.29	66.89
	(M)	66.50	66.51	68.11	68.12	54.65	51.53	64.78	61.69
Oesophagus	(F)	12.79	12.77	16.15	16.13	10.53	9.01	23.51	20.49
	(M)	7.48	7.43	7.94	7.89	9.70	8.72	24.12	21.98
Stomach	(F)	29.91	29.84	30.61	30.53	16.26	12.76	23.49	18.79
	(M)	27.11	27.06	27.29	27.24	16.23	13.93	27.13	23.69
Colon	(F)	57.55	57.38	57.78	57.61	48.49	41.31	59.40	51.98
	(M)	55.95	55.82	57.53	57.40	48.66	43.00	58.26	52.35
Rectum	(F)	55.71	55.66	55.91	55.87	53.24	48.75	58.40	53.86
	(M)	55.33	55.31	57.05	57.03	50.76	47.98	59.52	56.66
Lung	(F)	13.94	13.83	16.21	16.09	8.97	7.29	21.08	17.60
	(M)	8.47	8.42	10.14	10.08	8.49	7.04	17.86	15.07
Melanoma	(F)	88.26	88.24	88.43	88.41	86.06	82.56	90.90	88.12
	(M)	81.40	81.39	81.59	81.59	74.57	71.55	80.24	77.50
Breast	(F)	86.82	86.80	87.13	87.11	81.96	77.40	86.26	82.26
Cervix uteri	(F)	66.77	66.72	66.86	66.81	66.19	62.35	71.29	67.69
Corpus uteri	(F)	85.10	85.05	85.41	85.37	75.52	72.76	80.00	77.46
Ovary	(F)	55.57	55.54	55.57	55.54	37.17	32.36	44.10	38.92
Prostate	(M)	80.88	80.84	83.06	83.02	76.72	70.16	87.75	81.69

caution. However, in almost all cases the proportion which could be expected to survive for five or more years after diagnosis exceeds the proportion 'missing'. Moreover, in cases diagnosed at Thames during 1992 the proportion of registered patients dying within five years that have no mention of cancer on the death certificate is only 6%.

This study demonstrates the large effects that DCOs and incompleteness can have on calculated survival figures. It is important to take variations in DCO proportion and/or completeness into consideration when comparing survival estimates between different populations.

Conflict of interest statement

All authors declare that they have no conflicts of interest.

Acknowledgements

We thank the anonymous referees of an earlier draft of this paper for their helpful comments.

REFERENCES

1. Berrino F, Sant M, Verdecchia A, Capocaccia R, Hakulinen T, Estève J, editors. *Survival of cancer patients in Europe: the EUROCARE study.* Lyon: IARC Scientific Publications No. 132; 1995. International Agency for Research on Cancer.
2. Berrino F, Capocaccia R, Estève J, Gatta G, Hakulinen T, Micheli M, Sant M, Verdecchia A, editors. *Survival of cancer patients in Europe: the EUROCARE-2 study.* Lyon: IARC Scientific Publications No. 151; 1999. International Agency for Research on Cancer.
3. The EUROCARE Working Group. EUROCARE-3: Cancer survival in Europe at the end of the 20th century. *Annal Oncol* 2003;**14**(Suppl 5).
4. Brenner H. Limitations of the death certificate only index as a measure of incompleteness of cancer registration. *Br J Cancer* 1995;**72**:506–10.
5. Berrino F. The EUROCARE Study: strengths, limitations and perspectives of population-based, comparative survival studies. *Annal Oncol* 2003;**14**(Suppl. 5):v9–v13.
6. Pollock AM, Vickers N. The impact on colorectal cancer survival of cases registered by 'death certificate only': implications for national survival rates. *Br J Cancer* 1994;**70**:1229–31.
7. Pollock AM, Vickers N. Why are a quarter of all cancer deaths in south-east England registered by death certificate only? Factors related to death certificate only registrations in the Thames Cancer Registry between 1987 and 1989. *Br J Cancer* 1995;**71**:637–41.
8. Brenner H, Hakulinen T. Population-based monitoring of cancer patient survival in situations with imperfect completeness of cancer registration. *Br J Cancer* 2005;**92**: 576–9.
9. Bullard J, Coleman MP, Robinson D, Lutz J-M, Bell J, Peto J. Completeness of cancer registration: a new method for routine use. *Br J Cancer* 2000;**82**:1111–6.
10. Brenner H, Gefeller O. An alternative approach to monitoring cancer patient survival. *Cancer* 1996;**78**:2004–10.
11. Statacorp. Stata Statistical Software: Release 8.0. College Station; 2003. Stata Corporation.
12. Prior B, Woodman CBJ, Collins S. International differences in survival from colon cancer: more effective care versus less complete cancer registration. *Br J Surg* 1998;**85**: 101–4.