available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Model based hybrid analysis of cancer patient survival

Hermann Brenner[a,*], *Timo Hakulinen*[b]

[a]*Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Bergheimer Strasse 20, D-69115 Heidelberg, Germany*

[b]*Finnish Cancer Registry, Institute for Statistical and Epidemiological Cancer Research, Liisankatu 21 B, FIN-00170 Helsinki, Finland*

ARTICLE INFO

Article history:
Received 9 October 2006
Received in revised form
9 January 2007
Accepted 15 January 2007
Available online 27 February 2007

Keywords:
Epidemiologic methods
Models
Statistical
Neoplasms
Prognosis
Registries
Survival

ABSTRACT

In many cancer registries, registration of new cases is somewhat less up-to-date than mortality follow-up. In such situations, hybrid analysis, a combination of cohort and period analysis, rather than 'pure' period analysis has been proposed to derive up-to-date survival estimates. We evaluate application and adaptation of a modelling strategy that has recently been introduced to enhance precision of period survival estimates, to 'hybrid type of data'. Using data from the Finnish Cancer Registry, we show that modelling again strongly increases precision of survival estimates. Furthermore, special models adapted to the hybrid type of data are shown to provide even more precise and, in a clear majority of cases, also more valid predictions of survival of recently diagnosed patients than models ignoring the hybrid type of data. Finally, we show that model-based estimation of and testing for recent trends may give different answers if period rather than hybrid modelling is used for hybrid type of registry data. We conclude that modelling is useful for both hybrid and period analyses of cancer survival, but the different data structure needs to be taken into account in the set-up of models.

1. Introduction

Monitoring cancer patient survival is an important task of population-based cancer registries. Period analysis, a new method of survival analysis introduced some 10 years ago,[1] has been shown to provide more up-to-date cancer survival estimates than traditional methods of survival analysis.[2–6] The principle of period analysis has been described in detail elsewhere.[1,7] Briefly, it consists of restricting the analysis to the survival experience of cancer patients in some recent time period, which is achieved by left truncation of observations at the beginning of that time period (in addition to right censoring of observations at its end). In the application of period analysis, there is a tradeoff between up-to-dateness and precision of survival estimates: increasing up-to-dateness by restricting the analysis to a relatively short recent time period, such as the most recent calendar year (compared to, say, the most recent five calendar years) for which cancer registry data are available, goes along with a loss of precision. We recently proposed a model-based approach, in which much more precise period estimates of survival for the most recent single calendar year can be obtained through modelling of survival trends within a time window encompassing multiple recent years.[8]

'Pure' period analysis (including modelled period analysis) requires that complete registration of new cancer cases is as up-to-date as mortality follow-up. However, in many cancer registries, completion of registration of new cases for a given calendar year often lags behind by one or two years compared to mortality follow-up. For example, at a given point of time, the registry database may be complete with respect to recording incident cases up to the end of 2003, while mortality follow-up of registered cases may already be complete up to the end of 2004. Recently, a modified type of analysis, denoted

* *Corresponding author:* Tel.: +49 6221 548140; fax: +49 6221 548142.
E-mail address: h.brenner@dkfz-heidelberg.de (H. Brenner).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2007.01.015

'hybrid analysis', has been proposed to derive up-to-date survival estimates in such situations.[9] The aim of this paper is to evaluate possible modelling strategies for deriving up-to-date, accurate and precise estimates of cancer survival in the 'hybrid type of data structure'.

2. Patients and methods

2.1. Database

Our analysis is based on data from the nationwide Finnish Cancer Registry, covering a population of about 5 million people, which is well known for its high levels of completeness and data quality.[10] For this analysis, we used a database encompassing patients diagnosed within half a century from 1953 to 2002, with a follow-up with respect to vital status until the end of 2003. We included patients aged 15 or older with a first diagnosis with one of 20 common forms of cancer between 1953 and 2002.

2.2. Statistical analysis

Throughout this paper, we present relative rather than absolute survival rates, as the former are most commonly reported by population-based cancer registries. Relative survival rates reflect the probability of surviving the cancer of interest rather than the total survival probability,[11,12] taking expected deaths in the absence of cancer into account. For this analysis, the expected numbers of deaths were derived from age, gender and calendar period specific mortality figures of the general population of Finland according to the so-called Ederer II method.[13]

In a first step, actual 5-year relative survival of patients diagnosed in 1997 and followed through 2002 was compared with the most up-to-date estimates of 5-year relative survival that would potentially have been available with 'hybrid type' of registry data up to 1997 (the year of diagnosis of this cohort), i.e. registration of new cancer cases completed up to the end of 1996 only, and mortality follow-up completed up to the end of 1997. The following types of survival analyses were applied (see Figs. 1 and 2): First, a conventional ('unmodelled') period analysis for the year 1997 only (Fig. 1, right framed column), ignoring the fact that patients diagnosed in 1997 do not contribute to the database, as they would in a true period analysis for the year 1997. Second, a 'modelled period analysis', as recently proposed for a period type of data constellation,[8] by which a period estimate of 5-year relative survival for 1997 is estimated by trend modelling from a database including 5 one-year periods (1993, 1994, 1995, 1996 and 1997) (Fig. 1, 5 framed columns). Third, a conventional ('unmodelled') hybrid analysis for the year 1997 only (taking part of the survival experience during the 1st year following diagnosis from the 1996 cohort, Fig. 2, right framed area), as previously described.[9] Fourth, a 'modelled hybrid analysis', by which a hybrid estimate of 5-year relative survival for 1997 is estimated by trend modelling from a database including five one-year hybrid type databases (1993, 1994, 1995, 1996 and 1997) (Fig. 2, 5 framed areas). Simultaneous presentation of results obtained with the four methods allows for a comparison of modelled and unmodelled hybrid analyses on one hand, and of (modelled or unmodelled) hybrid and period analyses on the other hand in the hybrid data situation.

The modelling approach was carried out exactly as previously described in detail,[8] except for the necessary adaptation to the slightly different database used in hybrid analysis. Briefly, survival probabilities were modelled for each combination of one-year database (framed areas in Figs. 1 and 2) and year of follow-up within the specified time window. For

Years of	Years of Follow-up									
Diagnosis	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
1988	1	1-2	2-3	3-4	4-5	5				
1989		1	1-2	2-3	3-4	4-5	5			
1990			1	1-2	2-3	3-4	4-5	5		
1991				1	1-2	2-3	3-4	4-5	5	
1992					1	1-2	2-3	3-4	4-5	**5**
1993						1	1-2	2-3	3-4	**4-5**
1994							1	1-2	2-3	**3-4**
1995								1	1-2	**2-3**
1996									1	**1-2**

Fig. 1 – Years of diagnosis and years of follow-up included in the calculations of period estimates of 5-year relative survival of patients for the years 1993–1997. The numbers within the cells indicate the years following diagnosis.

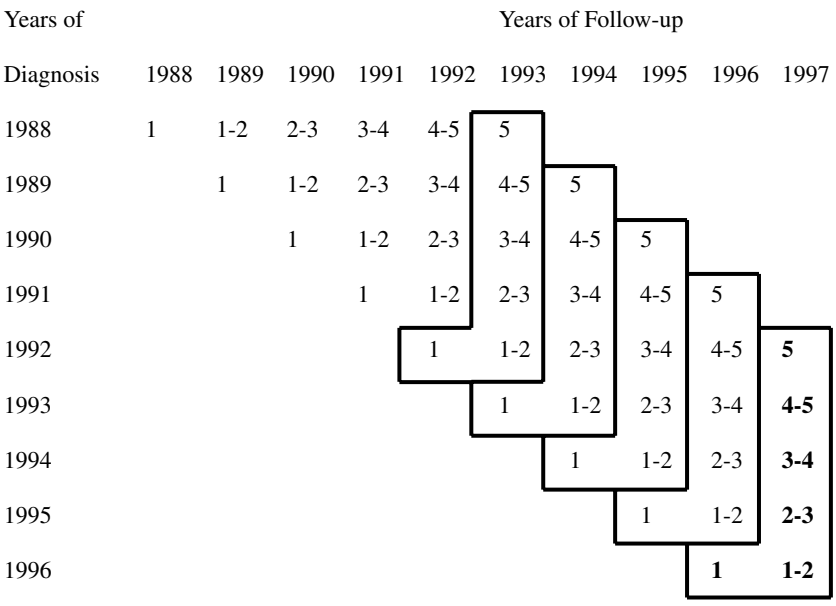

Fig. 2 – Years of diagnosis and years of follow-up included in the calculations of hybrid estimates of 5-year relative survival of patients for the years 1993–1997. The numbers within the cells indicate the years following diagnosis.

that purpose, numbers of patients at risk and of deaths by year of follow-up were first calculated for each single one-year database. Next, a Poisson regression model for relative survival was employed,[14] in which the numbers of deaths for each combination of one-year database and year of follow-up were modelled as a function of one-year database (included as a numerical predictor variable) and year of follow-up (included as a categorical predictor variable). The logarithm of the person-years at risk was used as offset, and late entries and withdrawals were accounted for as half-persons. Conditional survival probabilities for each year of follow-up, 5-year cumulative period survival estimates, and their standard errors were derived from the model results as previously described.[8]

The model based approach provides a general framework which encompasses conventional period or hybrid analyses as special cases of applications of saturated models. For example, a conventional period estimate of 5-year survival for 1997 can be obtained from a saturated model, in which the period of interest includes just one calendar year, i.e. 1997. This way, only five observations are included in the regression model, from which five parameters are estimated (one for each year of follow-up, none for calendar year). To ensure perfect comparability of results for conventional and modelled period and hybrid analyses, we derived the conventional estimates as special cases from saturated models by the same computer programs used for the modelling approaches in our analysis.

Next, we repeated the analyses explained for the year 1997 in the preceding paragraphs for each single calendar year from 1962 to 1997 (which is the widest possible range of calendar years for which pertinent calculations could be carried out with the aforementioned database of the Finnish Cancer Registry). This allowed to address the performance of the various methods in a much broader range of settings. We calculated the following summary indicators of the performance of the various methods: The mean difference and the mean squared difference between 5-year relative survival later observed for patients diagnosed in the respective year and the various estimates potentially available by the end of that year. The mean differences reflect the average under- or overestimation of the 5-year relative survival. The mean squared differences in addition reflect, among other factors, the random variation in the various estimates.

Finally, we employed period and hybrid modelling to estimate recent trends in 5-year survival in 1999–2003.

The analyses were carried out using the SAS statistical software package (Cary, NC). For all survival analyses, the macro *period* was used to derive the numbers of patients at risk and of deaths by year of follow-up and by calendar year.[7,15] Some minor formal modification of the output was made to facilitate the subsequent steps. Next, the procedure GENMOD was used to carry out Poisson regression, and the output of the regression models was used to carry out the subsequent calculations as previously described.[8]

3. Results

Overall, 682,867 patients aged 15 or older were reported to the Finnish Cancer Registry with a first diagnosis of cancer between 1953 and 2002. Of these, we excluded 2.3% notified by death certificate only, another 2.5% notified by autopsy only, and 0.1% due to missing information on month of diagnosis. The 20 forms of cancer specifically addressed in this paper include about 89.9% of the remaining cancer cases.

The numbers of notifications of patients with these 20 forms of cancer in 1997, as well as the proportions of female patients and mean age of patients, are shown in Table 1. Breast cancer was the most common form of cancer, followed by prostate cancer and lung cancer. Mean age at diagnosis varied from 52 years (thyroid cancer) to 72 years (liver cancer, prostate cancer).

Table 1 – Number, proportion female and mean age of patients diagnosed with common forms of cancer aged 15 and older in Finland in 1997

Cancer site	Number	Proportion female (%)	Mean age
Oral cavity	479	36	66
Oesophagus	206	45	70
Stomach	833	46	71
Colon	1211	57	69
Rectum	832	48	69
Liver	409	57	72
Pancreas	680	51	70
Lung	1977	23	69
Breast	3195	100	61
Cervix	137	100	60
Corpus	656	100	67
Ovaries	502	100	63
Prostate	2841	0	72
Kidneys	648	41	65
Urinary bladder	790	24	71
Melanoma	561	53	59
Brain	364	48	56
Thyroid gland	344	77	52
Leukaemias	360	43	66
Lymphomas	556	48	64

Actual 5-year relative survival later observed for patients diagnosed in 1997 varied from 90.6% (thyroid cancer) to 3.2% (pancreas cancer) (see Table 2). Based on incidence data up to 1996 and mortality data up to 1997, the estimates of 5-year relative survival for 1997 obtained by conventional (unmodelled) period analysis most often showed the largest difference from the later observed 5-year relative survival (for 9 out of 20 forms of cancer, underlined figures in Table 2). These estimates also had the largest standard errors. Application of conventional (unmodelled) hybrid analysis rather than conventional period analysis reduced the standard errors to some extent, and further substantial reduction of standard errors was achieved by application of modelled period analysis and particularly modelled hybrid analysis. The modelled period and hybrid estimates also most often came closest to the later observed 5-year relative survival (for 8 and 9 forms of cancer, respectively; bold figures in Table 2).

The unmodelled period analysis clearly performed worst in the evaluation including each single calendar year between 1962 and 1997 (see Table 3). The mean difference from the later observed 5-year relative survival of patients diagnosed in the respective years, and the mean squared difference was highest using this method for 13 and 17 out of 20 forms of cancer, respectively (underlined figures in Table 3). Although conventional (unmodelled) hybrid analysis performed somewhat better, the best results were obtained with the modelling approaches. In particular, hybrid modelling performed best with respect to the mean difference and the mean squared difference for 5 and 13 forms of cancer, respectively (bold figures in Table 3), whereas it performed worst with respect to these criteria for 1 and 0 forms of cancer only,

Table 2 – Various point estimates (PE) of 5-year relative survival potentially available with incidence data up to 1996 and follow-up data up to 1997 and of 5-year relative survival later observed for patients diagnosed in Finland in 1997

Cancer site	Observed		Period unmodelled		Hybrid unmodelled		Period modelled		Hybrid modelled	
	PE	SE	PE	SE	PE	SE	PE	SE	PE	SE
Best performance of hybrid modelling[a]										
Oral cavity	68.1	2.8	65.2	3.3	67.5	2.9	65.9	2.3	**66.0**	2.1
Rectum	55.7	2.1	58.2	2.6	54.4	2.3	56.6	1.8	**56.2**	1.7
Liver	6.5	1.5	14.3	2.7	14.6	2.2	12.0	1.4	**11.7**	1.2
Pancreas	3.2	0.8	**1.9**	0.6	4.2	1.1	**3.5**	0.6	**4.5**	0.6
Cervix	64.5	4.6	61.6	5.0	59.6	4.6	66.6	3.6	**64.9**	3.4
Corpus	81.5	2.0	83.4	2.2	82.7	2.0	83.4	1.5	**82.6**	1.5
Ovaries	45.1	2.4	39.9	2.8	39.3	2.4	**41.1**	2.0	**41.1**	1.8
Brain	32.3	2.5	27.0	3.1	29.2	2.4	28.7	2.1	**29.3**	1.9
Thyroid gland	90.6	2.0	96.3	1.6	91.6	2.0	92.7	1.4	**91.4**	1.4
Intermediate performance of hybrid modelling[a]										
Oesophagus	9.1	2.2	4.7	1.8	11.4	3.2	**8.5**	1.7	10.2	1.6
Colon	59.4	1.8	61.2	2.3	57.1	1.9	**59.3**	1.5	57.2	1.3
Lung	10.1	0.8	8.8	0.8	10.5	0.8	**10.2**	0.6	11.0	0.5
Prostate	82.2	1.3	**73.5**	1.7	73.0	1.6	**73.5**	1.2	73.4	1.2
Kidneys	59.7	2.3	61.1	2.7	56.6	2.3	**58.8**	1.9	56.9	1.7
Urinary bladder	70.6	2.4	73.2	2.7	**73.1**	2.4	75.4	1.9	74.9	1.7
Melanoma	83.5	2.1	82.0	2.4	82.6	2.2	**83.1**	1.7	82.6	1.7
Leukaemias	37.0	2.9	50.2	3.9	**43.6**	3.0	47.6	2.5	44.7	2.1
Lymphomas	50.2	2.4	**48.8**	2.9	45.6	2.4	48.4	2.0	47.8	1.8
Worst performance of hybrid modeling[a]										
Stomach	31.6	1.9	**29.3**	2.2	28.6	1.8	28.7	1.4	27.7	1.2
Breast	85.6	0.8	**84.0**	1.0	83.4	0.9	83.3	0.7	83.2	0.7

In addition, the standard errors (SE) of all estimates are given. Bold (underlined) figures indicate the estimates that come closest to (show largest difference from) the later observed 5-year relative survival.
a in terms of difference from later observed 5-year relative survival.

Table 3 – Mean difference and mean squared difference of the various estimates of 5-year relative survival of cancer patients in Finland potentially available with follow-up data up to each single year between 1962 and 1997 from 5-year relative survival later observed for patients diagnosed in those years[a]

Cancer site	Mean difference				Mean squared difference			
	Unmodelled		Modelled		Unmodelled		Modelled	
	Period	Hybrid	Period	Hybrid	Period	Hybrid	Period	Hybrid
Best performance of hybrid modelling[b]								
Oral cavity	−0.89	0.03	−0.37	**0.02**	16.95	16.32	14.89	**13.03**
Oesophagus	−4.14	**0.04**	−1.98	−0.21	25.25	8.38	10.18	**5.80**
Stomach	−3.45	**−1.12**	−1.91	−1.18	16.15	3.98	5.84	**3.61**
Liver	−3.89	−0.22	−1.22	**0.12**	24.36	9.75	9.72	**8.04**
Pancreas	−2.68	0.05	−1.10	**0.02**	9.19	3.26	3.12	**2.04**
Lung	−2.96	−0.11	−1.26	**−0.04**	10.38	1.39	2.57	**0.96**
Cervix	−0.83	**−0.43**	−0.65	−0.55	43.54	34.58	27.63	**27.22**
Corpus	0.82	−0.62	**0.07**	−0.55	11.93	10.85	7.47	**7.25**
Kidneys	**0.63**	−1.58	−0.68	−1.72	20.90	19.27	18.43	**18.12**
Urinary bladder	**−0.70**	−1.85	−1.16	−1.82	32.24	24.37	17.88	**16.99**
Melanoma	−2.09	−1.93	−1.84	**−1.78**	44.34	40.10	32.33	**32.08**
Thyroid gland	4.08	−1.44	**0.91**	−1.46	59.30	26.28	19.83	**17.77**
Leukaemias	3.70	−1.35	**0.45**	−1.55	42.78	17.32	14.67	**13.57**
Intermediate performance of hybrid modelling[b]								
Colon	3.00	−1.21	**0.69**	−1.08	23.37	**11.31**	11.34	11.40
Rectum	**0.19**	−1.57	−0.61	−1.44	**14.92**	16.52	15.72	15.46
Breast	**−1.21**	−1.76	−1.49	−1.70	**5.96**	7.95	7.06	7.73
Ovaries	0.97	**−0.10**	0.57	0.19	20.10	**13.48**	13.95	13.86
Prostate	**−1.16**	−2.60	−1.49	−2.19	21.83	24.62	**19.44**	20.70
Brain	**0.12**	−0.82	−0.35	−0.77	28.54	20.75	**19.48**	20.09
Lymphomas	**−0.66**	−1.21	−0.82	−1.00	34.42	23.44	**21.61**	21.93

Bold (underlined) figures indicate which method performs best (worst) according to each criterion.
a Availability of incidence data is assumed to lag one year behind availability of follow-up data throughout.
b in terms of mean squared difference.

respectively. In general, all approaches tended to provide slightly too pessimistic estimates, as the majority of mean differences shown in Table 3 were negative. However, in contrast to all other methods, substantial overestimation was also seen using the conventional period approach for some cancers, and underestimation exceeding three or more percent units on average was seen for some other cancers.

The modelled period and hybrid analyses for the 1999–2003 period (with additional follow-up data during the 1st year following diagnosis from 1998 included in the modelled hybrid analysis) indicated an increase in 5-year relative survival for most cancer sites, which was particularly pronounced for prostate cancer, lymphomas, breast and cervical cancer (see Table 4). Although findings from period and hybrid modellings were similar for most cancers, differences were not always negligible, and for some cancers (cervical cancer, leukaemias, thyroid cancer), significance of trends (at $\alpha = 0.05$) varied according to the type of modelling.

4. Discussion

Ten years after its introduction,[1] and a few years after its thorough evaluation by various groups,[2–6] the period analysis approach is meanwhile used by an increasing number of cancer registries to derive up-to-date estimates of cancer survival (e.g. Refs. [16–24]). However, in many cancer registries, the delayed availability of complete incidence data compared to mortality follow-up asked for modification of period analysis, leading to the concept of hybrid analysis.[9]

In this paper, we introduced and evaluated an extension of the recently proposed modelling approach for period survival analysis [8] to the hybrid type of data available in many cancer registries. We demonstrated that modelling may be as useful for hybrid type of data as for pure period type of data. Our results furthermore underline the importance to take the different data structure into account even in the modelling approach. In the hybrid data situation, application of hybrid type modelling provided both more precise and more valid predictions of 5-year relative survival of recently diagnosed patients for most forms of cancer than period modelling. Furthermore, we showed that estimation of recent trends by modelling may sometimes give different answers if period rather than hybrid modelling is used for hybrid type of registry data.

Our results concerning conventional (unmodelled) hybrid analysis are in agreement with previous findings,[9] and they underline the need to switch from pure period analysis to hybrid analysis in the hybrid data situation. A serious drawback of application of pure period analysis in that situation is the loss of precision due to the sparseness of data on survival experience early after diagnosis, which is caused by delayed availability of registration of recent incident cases. In addition, however, estimates may also be biased due to the under-representation of the survival experience in the early

Table 4 – Modelled period and hybrid estimates and p-values for trend of 5-year relative survival for cancer patients in Finland in 1999–2003

Cancer site	Modelled period estimates				Modelled hybrid estimates			
	1999	2003	Change	p-Value[a]	1999	2003	Change	p-Value[a]
Significant change in 5-year relative survival according to both modelled period and hybrid analyses								
Colon	58.2	63.9	5.7	0.01	57.2	61.3	4.1	0.03
Rectum	53.8	60.0	6.2	0.02	53.5	58.6	5.1	0.03
Lung	11.9	8.6	−3.3	<0.001	11.8	9.8	−2.0	0.003
Breast	83.6	90.5	6.9	<0.0001	83.4	90.2	6.8	<0.0001
Ovaries	45.0	52.3	7.3	0.03	44.2	49.9	5.7	0.04
Prostate	78.8	90.9	12.1	<0.0001	79.0	90.6	11.6	<0.0001
Lymphomas	49.4	59.8	10.4	<0.001	49.0	57.2	8.2	0.001
Significant change according to period modelling only								
Cervix	57.9	72.1	14.2	0.02	60.1	70.0	9.9	0.06
Leukaemias	37.0	46.7	9.7	0.01	38.2	42.2	4.0	0.18
Significant change according to hybrid modelling only								
Thyroid gland	88.6	92.7	4.1	0.14	86.6	91.9	5.3	0.02
No significant change								
Oral cavity	65.5	68.7	3.2	0.35	66.2	68.5	2.3	0.43
Oesophagus	10.4	10.7	0.3	0.91	11.1	11.6	0.5	0.79
Stomach	31.1	30.1	−1.0	0.67	31.0	30.7	−0.3	0.86
Liver	9.2	11.1	1.9	0.30	9.6	11.2	1.6	0.22
Pancreas	3.9	3.3	−0.6	0.45	3.8	4.3	0.5	0.35
Corpus	84.0	84.0	0.0	0.99	83.9	83.0	−0.9	0.66
Kidneys	58.7	62.2	3.5	0.24	58.5	59.3	0.8	0.75
Urinary bladder	70.3	74.2	3.9	0.19	70.8	72.4	1.6	0.52
Melanoma	84.1	84.6	0.5	0.88	84.6	84.0	−0.6	0.79
Brain	31.2	30.3	−0.9	0.79	31.6	31.0	−0.6	0.85

a Two-sided p-value for test of linear trend in 1999–2003.

part of the 1st year following diagnosis in the hybrid data situation. For example, in the period analysis for 1997 using the database shown in the right most column in Fig. 1, survival experience during the 1st year following diagnosis exclusively comes from patients diagnosed in 1996, whose survival experience immediately following diagnosis occurred in 1996 and remains unconsidered. This differential neglect of survival experience immediately following diagnosis is overcome by switching to hybrid analysis.

Disproportional representation of survival experience in the early and late part of the 1st year following diagnosis is of particular importance as, for many forms of cancer, fatality is relatively high and often not homogeneous during the 1st year following diagnosis. Although the differential representation of various parts of the 1st year following diagnosis is somewhat reduced overall in the 5-year period used for period modelling, the differential representation in the final year of that period persists and may give rise to an erroneous trend estimate if the hybrid type of data is neglected in the analysis. These considerations and the results of our empirical evaluation therefore underline the importance to apply a hybrid type of analysis even with modelling of hybrid type of registry data over a 5-year time frame.

In addition to benefits with respect to validity, hybrid modelling also increases precision of survival estimates compared with period modelling. This gain in precision is due to the larger database included in the modelling. In particular, person-time at risks and deaths during the first year following diagnosis (when a large proportion of cancer deaths occurs) are higher in the hybrid modelling approach.

In our evaluation, we only considered one special example of hybrid type of registry data, i.e. a situation where registration of incident cases lags one year behind mortality follow-up. Although this may be the most commonly encountered hybrid data situation in practice, a longer time lag, such as two years or even three years, may be encountered in some registries. In such situations, application of a hybrid type of analysis, which would only require a slight modification to the approach illustrated in this paper, may even be more important.

An alternative to the use of hybrid analysis would be to exclude the most recent years altogether, for which incidence data are not yet available, and to conduct a pure period analysis for the most recent year(s) for which such data are available. An obvious drawback of this approach would be an unnecessary loss of up-to-dateness of survival estimates, which may be overcome by application of the hybrid type of analysis.

In summary, we conclude that modelling is as useful for hybrid type of data as for pure period type of data, but the different data structure should be taken into account in the set-up of the models.

Conflict of interest statement

None.

Acknowledgements

The work of Hermann Brenner was partly supported by a grant from the German Cancer Foundation (Deutsche Krebshilfe, Project No. 70-3166-Br 5). Timo Hakulinen's work was supported by grants from the Academy of Finland and the Cancer Society of Finland.

REFERENCES

1. Brenner H, Gefeller O. An alternative approach to monitoring cancer patient survival. *Cancer* 1996;**78**:2004–10.
2. Brenner H, Hakulinen T. Advanced detection of time trends in long-term cancer patient survival: experience from 50 years of cancer registration in Finland. *Am J Epidemiol* 2002;**156**:566–77.
3. Brenner H, Hakulinen T. Up to date survival curves of patients with cancer by period analysis. *J Clin Oncol* 2002;**20**:826–32.
4. Brenner H, Söderman B, Hakulinen T. Use of period analysis for providing more up-to-date estimates of long-term survival rates: empirical evaluation among 370,000 cancer patients in Finland. *Int J Epidemiol* 2002;**31**:456–62.
5. Talbäck M, Stenbeck M, Rosén M. Up-to-date long-term survival of cancer patients: an evaluation of period analysis on Swedish Cancer Registry data. *Eur J Cancer* 2004;**40**:1361–72.
6. Ellison L. An empirical evaluation of period survival analysis using data from the Canadian Cancer Registry. *Ann Epidemiol* 2006;**16**:191–6.
7. Brenner H, Gefeller O, Hakulinen T. Period analysis for 'up-to-date' cancer survival data: theory, empirical evaluation, computational realisation and applications. *Eur J Cancer* 2004;**40**:326–35.
8. Brenner H, Hakulinen T. Up-to-date and precise estimates of cancer patient survival: model based period analysis. *Am J Epidemiol* 2006;**164**:689–96.
9. Brenner H, Rachet B. Hybrid analysis for up-to-date long-term survival rates in cancer registries with delayed recording of incident cases. *Eur J Cancer* 2004;**40**:2491–501.
10. Teppo L, Pukkala E, Lehtonen M. Data quality and quality control of a population-based cancer registry. Experience in Finland. *Acta Oncol* 1994;**33**:365–9.
11. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. *Monogr Natl Cancer Inst* 1961;**6**:101–21.
12. Henson DE, Ries LA. The relative survival rate. *Cancer* 1995;**76**:1687–8.
13. Ederer F, Heise H. Instructions to IBM 650 programmers in processing survival computations. Methodological note No. 10, End Results Section. Bethesda (MD), National Cancer Institute; 1959.
14. Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Stat Med* 2004;**23**:51–64.
15. Arndt V, Talbäck M, Gefeller O, Hakulinen T, Brenner H. Modification of SAS macros for more efficient analysis of relative survival rates. *Eur J Cancer* 2004;**40**:778–9.
16. Aareleid T, Brenner H. Trends in cancer patient survival in Estonia before and after the transition from a Soviet republic to an open market economy. *Int J Cancer* 2002;**102**:45–50.
17. Brenner H. Long-term survival rates of cancer patients achieved by the end of the 20th century: a period analysis. *Lancet* 2002;**360**:1131–5.
18. Coleman MP, Rachet B, Woods LM, et al. Trends and socioeconomic inequalities in cancer survival in England and Wales up to 2001. *Brit J Cancer* 2004;**90**:1367–73.
19. Talbäck M, Rosén M, Stenbeck M, Dickman PW. Cancer patient survival in Sweden at the beginning of the third millenium – predictions using period analysis. *Cancer Causes Control* 2004;**15**:967–76.
20. Brenner H, Stegmaier C, Ziegler H. Long-term survival of cancer patients in Germany achieved by the beginning of the 3rd millennium. *Ann Oncol* 2005;**16**:981–6.
21. Houterman S, Janssen-Heijnen ML, van de Poll-Franse LV, Brenner H, Coebergh JW. Higher long-term survival rates in southeastern Netherlands using up-to-date period analysis. *Ann Oncol* 2006;**17**:709–12.
22. Stang A, Valiukeviciene S, Aleknaviciene B, Kurtinaitis J. Time trends of incidence, mortality and relative survival of invasive skin melanoma in Lithuania. *Eur J Cancer* 2006;**42**:660–7.
23. Zuccolo L, Dama E, Maule MM, Pastore G, Merletti F, Magnani C. Updating long-term childhood cancer survival trend with period and mixed analysis: Good news from population-based estimates in Italy. *Eur J Cancer* 2006;**42**:1135–42.
24. Ellison LF, Gibbons L. Survival from cancer – up-to-date prediction using period analysis. *Health Rep* 2006;**17**:19–30.