

Up-to-Date Estimates of Cancer Patient Survival Even with Common Latency in Cancer Registration

Hermann Brenner¹ and Timo Hakulinen²

¹Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany; and
²Finnish Cancer Registry, Institute for Statistical and Epidemiological Cancer Research, Helsinki, Finland

Abstract

In an era of ongoing improvement in cancer patient survival, available long-term survival figures from cancer registries are often outdated and too pessimistic for two reasons: first, delay in availability of cancer registry data, typically in the order of a few years, and, second, application of cohort-based methods of survival analysis, which provide survival estimates for patients diagnosed many years ago. We developed a model-based period analysis approach aimed to overcome both problems. We provide extensive empirical evaluation of our approach by comparing its performance with that of previously available methods for monitoring of 5- and 10-year relative survival,

with the use of data from the nationwide Finnish Cancer Registry of 490,279 patients ages ≥ 15 years and diagnosed with one of 20 common forms of cancer between 1953 and 1997. We show that, in most cases, the model-based approach predicts 5- and 10-year relative survival expectations of newly diagnosed patients quite closely and much better than any of the previously available methods, including standard period analysis. We conclude that the model-based approach may enable deriving up-to-date cancer survival rates even with the common latency in availability of cancer registry data. (Cancer Epidemiol Biomarkers Prev 2006;15(9):1727–32)

Introduction

Monitoring cancer patient survival is an important task of both clinical and population-based cancer registries. To be of maximum use for both clinical and public health purposes, estimates of cancer patient survival should be as up-to-date as possible. Period analysis has been introduced 10 years ago (1) to provide more up-to-date estimates of long-term survival than traditional cohort-based methods of survival analysis. The principle of period analysis consists of restricting the analysis to 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). It has been shown that period analysis provides quite accurate and up-to-date estimates of long-term survival for patients diagnosed in the period of interest (2–5). In practice, however, there is often a latency of several years in availability of cancer registry data, and there may be further delay in the process of publication. Therefore, even if period analysis is applied to provide up-to-date estimates of long-term cancer patient survival for the most recent period for which cancer registry data are available, these estimates may still be somewhat outdated with respect to currently diagnosed cancer patients. For example, period estimates of long-term cancer survival published in 2002 for the United States and in 2005 for Germany pertained to patients diagnosed with cancer up to the year 1998 and 2002, respectively (6, 7).

In this article, we develop and empirically evaluate a model-based period analysis approach aimed to provide up-to-date estimates of long-term survival even with the common latency in availability of cancer registry data.

Materials and Methods

Database. Our analysis is based on data from the nationwide Finnish Cancer Registry, which is well known for its high levels of completeness and data quality (8). At the time of this analysis, the database encompassed patients diagnosed within half a century from 1953 to 2002, with a follow-up with respect to vital status until the end of 2003. In this analysis, we included patients ages ≥ 15 years with a first diagnosis with 1 of 20 common forms of cancer between 1953 and 1997.

Statistical Analysis. Throughout this article, we present relative rather than absolute survival rates, as the former are most commonly reported by cancer registries. Relative survival rates reflect the probability of surviving the cancer of interest rather than the total survival probability (9, 10), 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 (11).

We first assessed overall trends in survival during the past decades by looking at the development of 5-year relative survival for successive cohorts of patients diagnosed between 1953 to 1957 and 1993 to 1997.

Next, 5-year relative survival actually observed for patients diagnosed in 1993 to 1997 and followed through 2002 (Fig. 1, *bottom solid rectangular frame*) was compared with the most up-to-date estimates of 5-year survival that would potentially have been available somewhere in the middle of 1993 to 1997, the years of diagnosis of this cohort. Assuming that, with the common latency of cancer registration of about several years, available data might have included patients diagnosed up to the year 1992, the following survival estimates were derived: First, a cohort-based estimate for the cohort of patients diagnosed in 1983 to 1987 and followed over 5 years since then (Fig. 1, *top solid rectangular frame*). Second, a so-called complete estimate additionally included patients diagnosed in 1988 to 1992, although the latter could not have been under observation for 5 years by the end of 1992 and might be censored at that date unless they died or were lost to follow-up earlier (Fig. 1, *triangular frame*). Third, a standard period estimate for the 1988

Received 3/20/06; revised 6/14/06; accepted 7/5/06.

Grant support: German Cancer Foundation, Deutsche Krebshilfe, project no. 70-3166-Br 5 (H. Brenner), and Academy of Finland and the Cancer Society of Finland (T. Hakulinen).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Hermann Brenner, Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Bergheimer Strasse 20, D-69115 Heidelberg, Germany. Phone: 49-6221-548140; Fax: 49-6221-548142. E-mail: h.brenner@dkfz-heidelberg.de

Copyright © 2006 American Association for Cancer Research.

doi:10.1158/1055-9965.EPI-06-0167

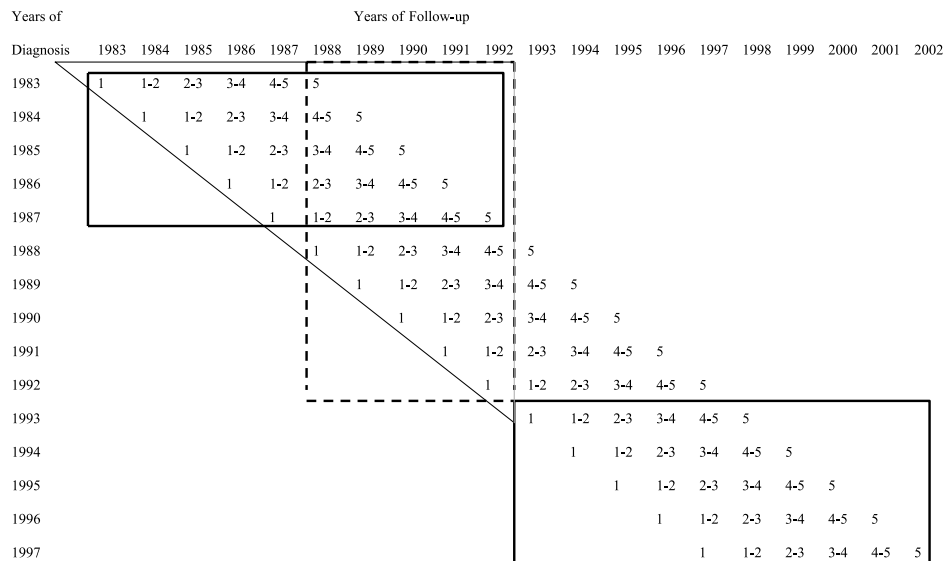


Figure 1. Years of diagnosis and years of follow-up included in calculations of 5-year relative survival of patients diagnosed in 1993 to 1997 (bottom solid rectangular frame) and in various period estimates of 5-year relative survival using data available in 1993 to 1997 (top solid rectangular frame: cohort estimate for the 1983 to 1987 cohort; triangular frame: complete estimate for calendar years 1983 to 1992; dashed rectangular frame: period estimate for the 1988 to 1992 period). Numbers within the cells, years following diagnosis.

to 1992 period, derived by left truncation of observations at the beginning of 1988 in addition to censoring at observations at the end of 1992 as previously described (Fig. 1, dashed frame; ref. 1). Fourth, to overcome the latency in cancer registration, a model-based period estimate for the 1993 to 1997 period was derived.

For derivation of the model-based period estimates for the 1993 to 1997 period, we first calculated numbers of patients at risk and of deaths by year of follow-up for each of the three preceding 5-year periods, i.e., 1978 to 1982, 1983 to 1987, and 1988 to 1992, just like one would do in standard period analysis for each of these periods. Next, we used a Poisson regression model (12) for the total 1978 to 1992 period. A formal description of the model is given in Appendix 1. Briefly, the numbers of deaths for each combination of 5-year calendar period and year of follow-up were modeled as a function of the calendar period (included as a numerical predictor variable) and year of follow-up (included as a categorical predictor variable). Based on this model, we calculated the

projected numbers of deaths and conditional relative survival probabilities for each year of follow-up in 1993 to 1997, assuming that the linear trend from the 1978 to 1982 period to the 1988 to 1992 period would prevail, and that the pattern of follow-up year-specific survival would remain unchanged otherwise. The model-based period estimate of 5-year relative survival for 1993 to 1997 was then obtained as the product of these conditional survival probabilities.

To address the performance of the various methods in a much broader range of settings, we repeated the analyses of 5-year relative survival outlined for the cohort of patients diagnosed in 1993 to 1997 in the preceding paragraphs for cohorts of patients diagnosed in 1988 to 1992, 1983 to 1987, 1978 to 1982, and 1973 to 1977 as well (which is the widest possible range of cohorts for which analogous calculations could be carried out with the available database of the Finnish Cancer Registry). We calculated the following summary indicators of the performance of the various methods: The

Table 1. Numbers and 5-year relative survival of patients diagnosed with common forms of cancer in Finland in 1953 to 1957 and 1993 to 1997

Cancer site	No. patients		5-y relative survival (%)			
	1953-1957	1993-1997	1953-1957	1993-1997	Difference	Mean annual difference
Oral cavity	1,428	2,311	66.3	66.4	0.1	0.00
Esophagus	1,305	1,000	3.2	10.1	6.9	0.17
Stomach	8,164	4,550	8.7	28.4	19.7	0.49
Colon	1,072	5,877	22.6	56.2	33.6	0.84
Rectum	891	3,897	23.8	54.4	30.6	0.77
Liver	396	2,251	4.7	10.5	5.8	0.14
Pancreas	839	3,257	5.0	3.7	-1.3	-0.03
Lung	4,456	10,144	6.4	10.2	3.8	0.09
Breast	3,131	15,094	52.5	83.9	31.3	0.78
Cervix	1,703	717	51.0	62.3	11.3	0.28
Corpus	969	3,135	63.0	82.4	19.4	0.48
Ovaries	751	2,379	25.3	41.4	16.1	0.40
Prostate	976	11,854	26.7	76.1	49.4	1.23
Kidneys	467	3,323	18.7	57.2	38.5	0.96
Urinary bladder	570	3,919	26.2	71.5	45.4	1.13
Melanoma	322	2,732	44.3	82.5	38.2	0.95
Brain	437	1,688	18.0	32.4	14.3	0.36
Thyroid gland	299	1,714	49.1	89.4	40.3	1.01
Leukemias	714	1,925	11.2	39.1	27.9	0.70
Lymphomas	338	2,725	22.0	49.0	27.0	0.68

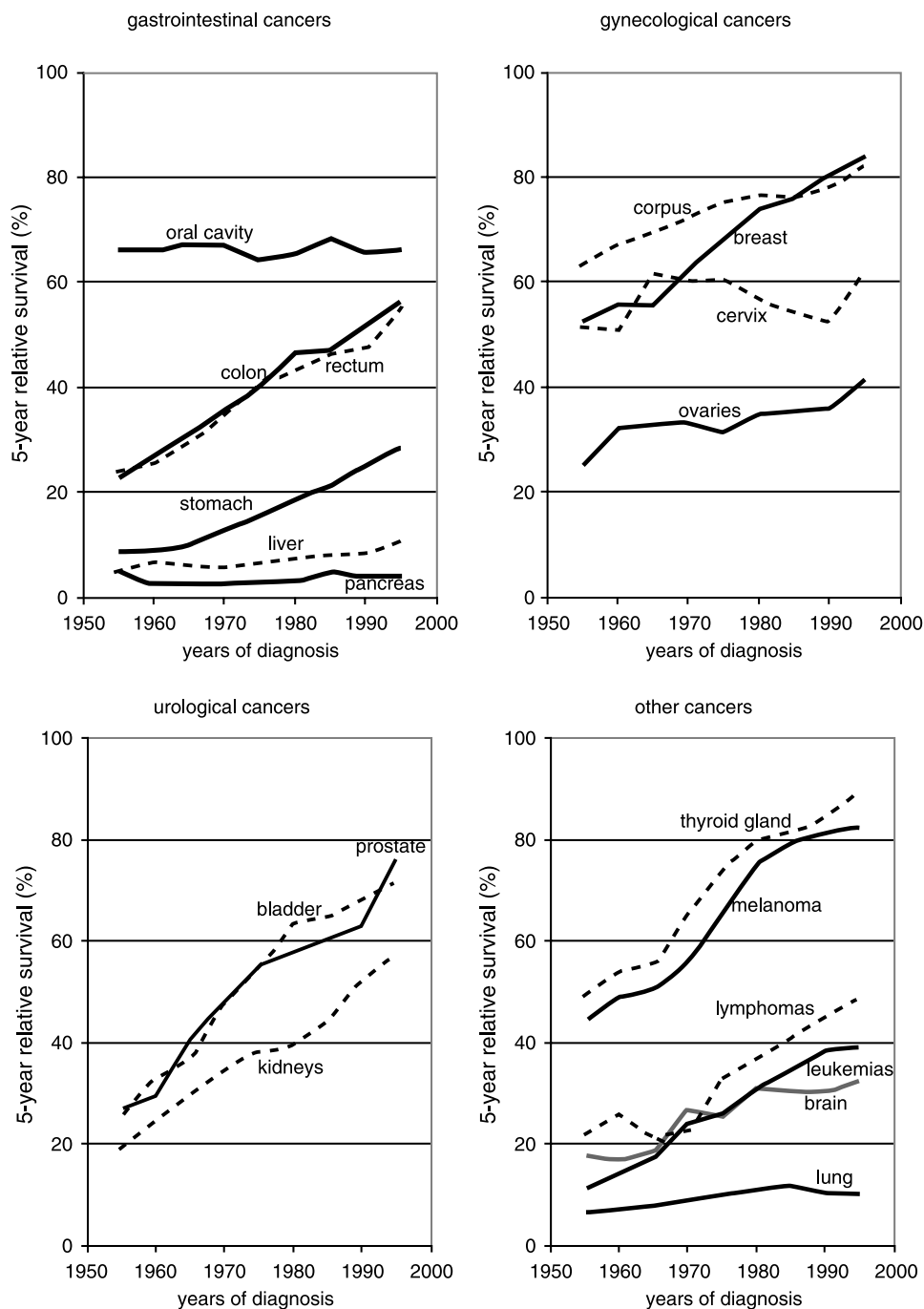


Figure 2. Trends in 5-year relative survival rates in the second half of the 20th century for common forms of cancer in Finland.

mean difference and the mean squared difference between 5-year relative survival later observed for patients diagnosed in the respective calendar periods and the various estimates potentially available during these periods. The mean differences reflect the average underestimation or overestimation of the 5-year relative survival rates. The mean squared differences, in addition, reflect average absolute levels of deviations of single estimates (with particular "punishment" of strong deviations).

To address the performance of the modeling approach for longer-term survival, we carried out analogous analyses comparing 10-year relative survival actually observed for patients diagnosed in 1988 to 1992, 1983 to 1987, and 1978 to 1982 with the most up-to-date estimates of 10-year relative survival that might have been available in those periods.

The analyses were carried out with the SAS statistical software package (Cary, NC), using the macro *period*, to derive

the effective numbers of patients at risk and of deaths (13), and the procedure GENMOD to carry out Poisson regression.

Results

Overall, 577,924 patients ages ≥ 15 years were reported to the Finnish Cancer Registry with a first diagnosis of cancer between 1953 and 1997. Of these, we excluded 2.6% notified by death certificate only, another 2.6% notified by autopsy only, and 0.1% due to missing information on month of diagnosis. The 20 forms of cancer specifically addressed in this article include $\sim 89.6\%$ of the remaining cancer cases ($n = 490,279$).

The numbers of notifications, as well as 5-year relative survival rates of patients with these 20 forms of cancer, are shown for calendar years 1953 to 1957 and 1993 to 1997 (Table 1). Five-year relative survival strongly varied by cancer

Table 2. Various estimates of 5-year relative survival potentially available in 1993 to 1997 compared with 5-year relative survival later observed for patients diagnosed in 1993 to 1997

Cancer site	5-y relative survival later observed for patients diagnosed in 1993-1997	Estimates of 5-y relative survival potentially available in 1993-1997 based on incidence and follow-up data up to the end of 1992			
		Cohort analysis	Complete analysis	Period analysis	
				Standard	Modeled
Oral cavity	66.4	<u>68.5</u>	67.2	65.9	<u>68.5</u>
Esophagus	10.1	<u>9.8</u>	<u>10.6</u>	10.4	<u>10.6</u>
Stomach	28.4	<u>21.0</u>	<u>22.5</u>	24.1	<u>26.7</u>
Colon	56.2	<u>47.4</u>	48.5	49.8	<u>52.8</u>
Rectum	54.4	<u>46.1</u>	<u>45.9</u>	47.4	<u>47.8</u>
Liver	10.5	<u>7.8</u>	8.6	9.0	9.1
Pancreas	3.7	<u>4.8</u>	4.6	4.6	4.1
Lung	10.2	<u>12.1</u>	11.2	10.6	9.6
Breast	83.9	<u>76.1</u>	77.3	78.5	80.9
Cervix	62.3	<u>54.4</u>	52.3	50.2	<u>45.6</u>
Corpus	82.4	<u>76.2</u>	76.7	77.2	<u>77.6</u>
Ovaries	41.4	<u>35.2</u>	<u>35.0</u>	35.6	38.7
Prostate	76.1	<u>60.4</u>	61.2	62.7	65.4
Kidneys	57.2	<u>44.2</u>	46.8	49.1	<u>54.2</u>
Urinary bladder	71.5	<u>64.8</u>	65.0	65.5	67.8
Melanoma	82.5	<u>79.4</u>	80.3	81.5	85.4
Brain	32.4	<u>30.4</u>	30.1	30.1	<u>29.6</u>
Thyroid gland	89.4	<u>81.6</u>	82.0	83.2	85.9
Leukemias	39.1	<u>34.5</u>	35.7	37.2	38.3
Lymphomas	49.0	<u>41.0</u>	41.2	41.9	45.8

NOTE: Bold (underlined) figures indicate the estimates that come closest to (show largest difference from) the later observed 5-year relative survival.

site. Among patients diagnosed in 1953 to 1957, it ranged from 66.3% for patients with cancer in the oral cavity to 3.2% for patients with esophagus cancer. Five-year relative survival rates increased between 1953 to 1957 and 1993 to 1997, albeit to a strongly varying degree, for all but one (pancreas cancer) of the assessed forms of cancer. A most pronounced increase by 49.4%, 45.4%, and 40.3% units (i.e., an average annual increase by >1% unit) was seen for patients with prostate, bladder, and thyroid cancer, respectively. With 89.4%, the latter had the highest 5-year relative survival rates in 1993 to 1997 among the cancer patients included in this analysis. On the other hand, 5-year relative survival still remained just above 10% for patients with cancers of the esophagus, liver, and lung, and <4% for patients with cancer of the pancreas.

As Fig. 2 shows, increases in 5-year relative survival mostly occurred steadily throughout the second half of the 20th century, and for some forms of cancer, such as colon, rectum, or stomach cancer, increases were very close to linear over several successive decades. On the other hand, quite varying trends were seen in different periods for a few other forms of cancer, such as lymphomas and particularly cervical cancer.

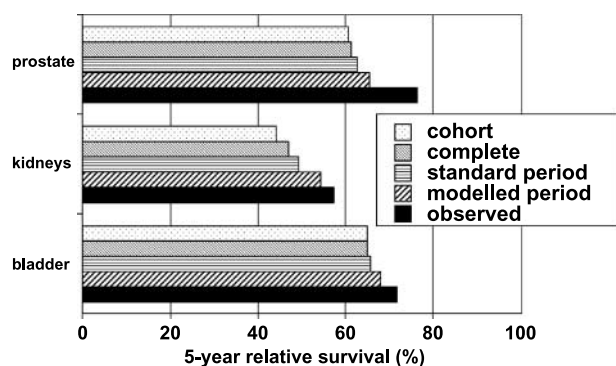


Figure 3. Various estimates of 5-year relative survival of patients with common urological cancers potentially available in 1993 to 1997 (based on incidence and follow-up data up to the end of 1992) compared with 5-year relative survival later observed for patients diagnosed in 1993 to 1997.

Table 2 shows the estimates of 5-year relative survival potentially available in 1993 to 1997 with cancer registry data up to the year 1992 by the four analytic approaches compared with the 5-year relative survival rates later observed for patients diagnosed in 1993 to 1997. For urological cancers, results are also illustrated graphically (see Fig. 3). For those forms of cancer whose prognosis improved over time, the available cohort estimates, which pertained to patients diagnosed in 1982 to 1987, but also (albeit to a slightly lesser degree) the available complete estimates, which pertained to patients diagnosed in 1982 to 1992, were much lower than the 5-year relative survival rates later observed in 1993 to 1997. The "standard" period estimates for the 1988 to 1992 period were less outdated in most cases, but for 14 of the 20 forms of cancer, the model-based period estimate for the 1993 to 1997 period came closest to the 5-year relative survival rates later observed in 1993 to 1997 (bold figures in Table 2). The latter was true for three, zero, and four forms of cancer for cohort analysis (which performed worst for 15 of 20 forms of cancer), complete analysis, and standard period analysis, respectively. However, for prostate and cervical cancer, all estimates were much too low. For prostate cancer, underestimation was worst for cohort analysis, and for cervical cancer, it was worst for the modeled period analysis.

With few exceptions, cohort, complete, and standard period analysis provided, on average, somewhat too pessimistic estimates of 5-year relative survival rates later observed for patients diagnosed in each 5-year calendar period between 1973 to 1977 and 1993 to 1997, which can be seen from the negative values of most of the mean differences shown in Table 3. This underestimation was generally largest for the conventional cohort analysis (with mean differences ranging up to -11.5% units for melanoma), intermediate for complete analysis, and somewhat lower for standard period analysis. With modeled period analysis, mean differences were generally much closer to 0. They were below 0 for 12 forms of cancer, and above 0 for eight forms of cancer, with a range from -3.5% units to +1.9% units. For 17 and 15 of 20 cancers, absolute values of mean differences and mean-squared differences were lowest with the use of modeled period analysis, respectively. Standard period analysis performed best for three and two forms of cancer, and cohort analysis

Table 3. Mean difference and mean squared difference of various estimates of 5-year relative survival available in calendar intervals 1973 to 1977, 1978 to 1982, 1983 to 1987, 1988 to 1992, and 1993 to 1997, and 5-year relative survival later observed for patients diagnosed in those calendar intervals

Cancer site	Mean difference				Mean squared difference			
	Cohort	Complete	Period, standard	Period, modeled	Cohort	Complete	Period, standard	Period, modeled
Oral cavity	0.4	0.4	0.4	<u>1.0</u>	6.8	6.4	7.7	<u>9.5</u>
Esophagus	<u>-1.3</u>	-0.8	-0.5	0.4	<u>2.5</u>	2.0	2.4	<u>2.2</u>
Stomach	<u>-6.2</u>	-5.1	-4.1	-2.0	<u>38.8</u>	26.6	16.9	<u>3.9</u>
Colon	<u>-8.3</u>	-6.7	-5.0	0.4	<u>72.4</u>	47.4	27.0	<u>5.5</u>
Rectum	<u>-7.6</u>	-6.7	-5.3	-0.8	<u>64.3</u>	50.2	31.8	<u>13.2</u>
Liver	<u>-1.5</u>	-0.8	-0.7	0.1	<u>2.6</u>	1.1	1.1	<u>0.8</u>
Pancreas	<u>-0.5</u>	-0.3	-0.4	-0.2	<u>1.5</u>	1.1	1.2	<u>0.8</u>
Lung	<u>-0.7</u>	-0.5	-0.4	<u>1.1</u>	<u>2.9</u>	1.4	1.1	<u>2.2</u>
Breast	<u>-9.3</u>	-7.8	-6.6	-2.6	<u>92.6</u>	65.2	47.0	<u>13.3</u>
Cervix	<u>1.5</u>	0.6	0.5	0.3	<u>26.2</u>	30.2	40.5	<u>79.8</u>
Corpus	<u>-3.7</u>	-3.3	-2.6	-0.8	<u>18.9</u>	14.4	10.3	<u>5.6</u>
Ovaries	<u>-2.3</u>	-1.9	-1.7	0.4	<u>12.3</u>	11.2	9.0	<u>8.7</u>
Prostate	<u>-10.1</u>	-8.6	-7.0	-0.2	<u>119.5</u>	89.2	62.4	<u>31.9</u>
Kidneys	<u>-9.0</u>	-7.2	-5.8	-1.3	<u>93.1</u>	60.4	40.5	<u>19.4</u>
Urinary bladder	<u>-11.0</u>	-9.1	-7.0	-0.2	<u>142.9</u>	94.6	57.0	<u>5.5</u>
Melanoma	<u>-11.5</u>	-10.1	-8.2	-3.5	<u>166.4</u>	131.6	96.5	<u>58.1</u>
Brain	<u>-3.4</u>	-2.3	-1.4	1.9	<u>17.7</u>	7.5	4.8	<u>15.1</u>
Thyroid gland	<u>-10.5</u>	-8.2	-6.2	-2.3	<u>136.8</u>	81.7	48.7	<u>23.6</u>
Leukemias	<u>-7.3</u>	-6.2	-4.5	-0.4	<u>54.8</u>	39.9	23.7	<u>1.2</u>
Lymphomas	<u>-10.1</u>	-8.7	-7.2	-2.9	<u>108.4</u>	79.3	56.1	<u>38.1</u>

NOTE: Bold (underlined) figures indicate the method that performed best (worst) according to each criterion.

performed worst for 18 and 18 forms of cancer, respectively, according to these two criteria. The only major exception from the good performance of modeled period analysis was cervical cancer, for which the mean squared difference was substantially higher with modeled period analysis than with the other methods of analysis.

The advantages of modeled period analysis are even more striking for 10-year relative survival rates, where underestimation of survival by traditional methods of survival analysis may be even much more severe, whereas the performance of the modeled period analysis is, on average, about equally good as for 5-year relative survival (see Table 4).

Discussion

With ongoing progress in prognosis, conventional cohort and complete estimates of long-term survival are overly pessimistic

for many forms of cancer. Although this problem may be partly overcome by period analysis, even the "standard" period estimates are often outdated to some extent once they can be derived due to the usual delay in availability of cancer registry data. In this article, we introduced a model-based extension of period analysis of cancer patient survival that allows to provide up-to-date estimates of long-term survival even in that situation. The model-based period analysis was found to be superior to any of the other methods, including standard period analysis, for the clear majority of cancers in extensive empirical evaluation.

Our findings regarding the advantage of "standard" period analysis over cohort and complete analysis in terms of up-to-dateness of survival estimates are in agreement with previous evaluations that had not taken the usual delay in availability of cancer registry data into account (3-5). Previous analyses had also shown that period analysis may advance

Table 4. Mean difference and mean squared difference of various estimates of 10-year relative survival available in calendar intervals 1978 to 1982, 1983-1987, and 1988 to 1992, and 10-year relative survival later observed for patients diagnosed in those calendar intervals

Cancer site	Mean difference				Mean squared difference			
	Cohort	Complete	Period, standard	Period, modeled	Cohort	Complete	Period, standard	Period, modeled
Oral cavity	1.5	0.6	1.6	<u>1.9</u>	3.4	4.9	16.0	<u>21.5</u>
Esophagus	<u>-2.2</u>	-1.8	-1.0	0.1	5.2	5.1	<u>6.7</u>	<u>5.8</u>
Stomach	<u>-8.7</u>	-6.6	-4.7	-2.7	<u>77.5</u>	45.2	22.6	<u>7.9</u>
Colon	<u>-12.7</u>	-8.5	-5.2	1.4	<u>161.3</u>	72.6	28.1	<u>4.1</u>
Rectum	<u>-9.6</u>	-6.3	-3.1	2.6	<u>104.3</u>	46.0	14.8	<u>10.9</u>
Liver	<u>-1.6</u>	-0.8	-0.4	0.1	<u>2.8</u>	0.8	0.7	<u>0.3</u>
Pancreas	<u>-0.6</u>	-0.3	-0.2	-0.1	<u>0.5</u>	0.3	0.4	<u>0.3</u>
Lung	<u>-0.9</u>	-0.5	-0.1	<u>1.6</u>	1.1	0.5	0.4	<u>3.6</u>
Breast	<u>-15.6</u>	-11.6	-8.7	-2.4	<u>242.6</u>	135.7	82.2	<u>14.0</u>
Cervix	8.7	7.4	6.5	6.8	75.8	57.1	50.0	<u>83.5</u>
Corpus	<u>-4.7</u>	-2.9	-1.3	0.7	<u>32.1</u>	15.7	9.7	<u>6.1</u>
Ovaries	<u>-1.3</u>	-1.1	-0.5	1.1	<u>6.3</u>	2.6	0.6	<u>4.1</u>
Prostate	<u>-11.0</u>	-7.9	-5.7	2.8	<u>127.9</u>	68.4	41.7	<u>17.2</u>
Kidneys	<u>-10.7</u>	-8.1	-6.1	-2.9	<u>124.6</u>	76.7	46.8	<u>19.7</u>
Urinary bladder	<u>-20.4</u>	-14.0	-9.3	-0.4	<u>433.5</u>	214.1	107.7	<u>9.0</u>
Melanoma	<u>-25.0</u>	-17.6	-11.5	-4.4	<u>640.4</u>	339.0	160.0	<u>58.1</u>
Brain	<u>-6.9</u>	-4.8	-3.5	0.7	<u>49.7</u>	24.3	18.6	<u>23.0</u>
Thyroid gland	<u>-18.0</u>	-10.9	-6.5	-0.6	<u>364.8</u>	127.1	42.6	<u>6.4</u>
Leukemias	<u>-8.5</u>	-6.4	-4.4	-0.1	<u>76.7</u>	44.6	24.7	<u>6.1</u>
Lymphomas	<u>-12.2</u>	-8.8	-5.9	1.6	<u>150.7</u>	77.5	37.9	<u>8.5</u>

NOTE: Bold (underlined) figures indicate the method that performed best (worst) according to each criterion.

detection of trends in 5-, 10-, 15-, and 20-year survival by almost 5, 10, 15, and 20 years compared with cohort analysis (2). Model-based period analysis as applied in this study would be expected to result in further reduction of delay in disclosure of progress in survival by another 5 years. In fact, the mean difference in model-based period estimates for the current period from 5-year and 10-year survival later observed for patients diagnosed in that period was close to zero for most forms of cancer in our extensive empirical evaluation, which indicates that the delay in disclosure of progress in prognosis can be overcome almost entirely.

In our modeling approach, we assumed persistent increasing or decreasing trends in relative survival rates over time. Obviously, the model-based period approach performs best when this assumption holds entirely. In our time trend analyses of cancer survival over half a century, we found entirely or close to monotonic upward trends in survival for most forms of cancer. However, for cervical cancer, divergent trends in various time intervals were found. The transient drop in survival for this form of cancer probably reflects selection processes resulting from the very successful screening program for this form of cancer (14). As a result of this pattern, the performance of model-based period estimates was worse for this form of cancer than the performance of the other methods of survival analysis. For cancers with little change in prognosis over time, neither standard period analysis nor model-based period analysis provides advantages over conventional cohort or complete analysis, and all methods essentially perform equally well. Such a pattern was seen, for example, for pancreas cancer in our analysis.

In the application of the modeling approach, there are a number of design options that may be considered. These include, for example, the number and width of periods included in the modeling. In our analysis, we chose to use three 5-year periods, covering a 15-year time frame altogether. Although a longer time frame (e.g., use of four or five 5-year periods) would provide a broader basis for estimation, the assumption of a persistent trend within such a long time frame may often be more problematic. Also, such an approach would only be feasible in cancer registries with a long history of cancer registration at high levels of quality and completeness. Although this prerequisite would be given in the Finnish Cancer Registry, it would hinder application of modeled period analysis in a large and growing number of younger high-quality cancer registries. On the other hand, relying on just two 5-year periods might provide a too weak basis for reliable estimation of trends and may give too much weight to possible random deviations from long-term trends. The width of time periods used for modeling and prediction is likewise somewhat arbitrary. We feel, however, that the 5-year periods used in our analysis may be a reasonable choice, as it limits the influence of short-term fluctuations in prognosis (whether they are due to chance or due to other reasons), and avoids the limited flexibility that would result from too broad intervals.

A potential tradeoff of the modeling approach, apart from its reliance on the existence of a long-term history of reliable cancer registration, is the increased complexity in calculation and data interpretation. However, the increased complexity of calculations also comes along with increased flexibility, e.g., to take account of possible deviations from linearity in trends, or of other variables that might affect prognosis that might be included in the models. In fact, the modeling approach may be considered as a broader general methodologic framework, which includes standard period analysis as one special case (i.e., the case of a saturated model in which follow-up year-specific survival rates are estimated for one single calendar period).

In summary, our empirical evaluation supports expectations from theory that the modeling approach further

enhances possibilities of deriving up-to-date cancer survival rates. This approach may be particularly helpful to overcome outdatedness of survival data resulting from the common latency in availability of cancer registry data.

Appendix 1

Let l_{ij} = the effective numbers of patients at risk (accounting for late entries and withdrawals as half persons), d_{ij} = the observed numbers of deaths, and e_{ij} = the expected numbers of deaths (from population life tables) for each combination of follow-up-year i and calendar period j .

The calendar periods are coded in such a way that $j = 0$ for the first calendar period, and $j = 1$ and 2 for the subsequent calendar periods included in the modeling.

Then, a generalized linear model $d_{ij} = f(i, j)$ is fitted with outcome d_{ij} , Poisson error structure, predictor variables i (categorical) and j (numerical), link $\ln(\mu_{ij} - d_{ij}^*)$, and offset $\ln(l_{ij} - d_{ij} / 2)$, where μ_{ij} = the model-based numbers of deaths and $d_{ij}^* = -(l_{ij} - d_{ij} / 2) \times \ln((l_{ij} - e_{ij}) / l_{ij})$.

Let α_i and β_j be the estimated regression coefficients for follow-up years i and calendar periods j . Then, estimates of conditional relative survival for each combination of follow-up year i and calendar period j are given as

$$r_{ij} = \exp[-\exp(\alpha_i + j \times \beta)],$$

and an estimate of cumulative k -year relative survival for each calendar period j is given as

$$R_j = \prod_{i=1}^k r_{ij} = \prod_{i=1}^k \exp[-\exp(\alpha_i + j \times \beta)]$$

The modeled k -year cumulative relative survival for the current calendar period is obtained by setting $j = 3$.

References

- Brenner H, Gefeller O. An alternative approach to monitoring cancer patient survival. *Cancer* 1996;78:2004–10.
- 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.
- Brenner H, Hakulinen T. Up to date survival curves of patients with cancer by period analysis. *J Clin Oncol* 2002;20:826–32.
- 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.
- Talback 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.
- 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.
- 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.
- 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.
- Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. *Monogr Natl Cancer Inst* 1961;6:101–21.
- Henson DE, Ries LA. The relative survival rate. *Cancer* 1995;76:1687–8.
- Ederer F, Heise H. Instructions to IBM 650 programmers in processing survival computations. Methodological note no. 10, end results section. Bethesda (Maryland): National Cancer Institute; 1959.
- Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Stat Med* 2004;23:51–64.
- Arndt V, Talback 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.
- Läära E, Day N, Hakama M. Trends in mortality from cervical cancer in the Nordic countries: association with organised screening programmes. *Lancet* 1987;1:1247–9.