American Journal of Epidemiology
Copyright © 2005 by the Johns Hopkins Bloomberg School of Public Health
All rights reserved
Vol. 161, No. 8
Printed in U.S.A.
DOI: 10.1093/aje/kwi099

PRACTICE OF EPIDEMIOLOGY

Substantial Overestimation of Standard Errors of Relative Survival Rates of Cancer Patients

Hermann Brenner[1] and Timo Hakulinen[2]

[1] German Centre for Research on Ageing, Heidelberg, Germany.
[2] Finnish Cancer Registry, Helsinki, Finland.

Received for publication September 27, 2004; accepted for publication December 9, 2004.

Relative survival rates are among the most commonly reported outcome measures of cancer patients. They are calculated as ratios of observed survival rates and the expected survival rates in the absence of cancer. Standard errors of relative survival rates are commonly calculated by dividing the standard error for absolute survival rates by the expected survival, without taking possible random variation of the latter into account. The aim of this study was to empirically assess the validity of these commonly reported standard errors. Using data from the nationwide Finnish Cancer Registry, the authors calculated 5- and 10-year absolute, expected, and relative survival rates for patients with 25 common forms of cancer in Finland in 1989. The authors used bootstrap analysis to empirically assess the random error of absolute and relative survival rates and then compared the results with conventionally derived estimates of standard errors. The conventional and bootstrap standard errors were closely similar for all estimates of absolute survival. By contrast, the conventional estimates of standard errors of 5- and 10-year relative survival exceeded the bootstrap estimates by up to 17% and 32%, respectively. The authors conclude that conventional derivation may substantially overestimate standard errors for relative survival.

epidemiologic methods; registries; survival

Relative survival rates are commonly reported by cancer registries. They are calculated as ratios of observed survival rates and the expected survival rates in the absence of cancer (1). The latter are typically derived from life tables of the general population. Standard errors of relative survival rates are commonly calculated by dividing the standard error for absolute survival rates by the expected survival, without taking possible random variation of the latter into account (2–4). The rationale for this procedure is that random variation in the expected survival can be neglected, as the database underlying the life tables from which it is derived is usually very large. This type of reasoning neglects the fact, however, that expected survival of a sample of patients is still subject to random variation due to random variation of the age distribution of the sample even if random variation of the life table estimates is negligible. Moreover, random errors of absolute and expected survival may often be positively correlated, because the age distribution of a sample of cancer patients may affect estimates of absolute and expected survival in a similar manner, as both absolute and expected survival are typically decreasing with increasing age. The aim of this study was to empirically assess the random error of expected survival rates resulting from random variation in the age distribution of cancer patients, its correlation with the random error of absolute survival rates, and the random error of relative survival rates, as well as to compare the results with estimates of standard errors obtained by neglecting random variation in expected survival rates altogether.

MATERIALS AND METHODS

This analysis is based on data from the nationwide Finnish Cancer Registry. This population-based registry has been operating for more than 50 years, covers the whole population of Finland (about 5.3 million people), and is

Correspondence to Dr. Hermann Brenner, Department of Epidemiology, German Centre for Research on Ageing, Bergheimer Str. 20, D-69115 Heidelberg, Germany (e-mail: brenner@dzfa.uni-heidelberg.de).

TABLE 1. Numbers and age distribution of patients aged 15 or more years with a first diagnosis of common forms of cancer, Finland, 1989

	No.	Age distribution (years)				
		10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
Oral cavity	416	46	58	67	77	83
Esophagus	200	57.5	67	73.5	81	87
Stomach	1,039	54	63	72	79	85
Small intestine	71	37	53	63	78	82
Colon	970	50	62	72	79	84
Rectum	599	54	62	70	78	83
Liver	430	57.5	65	72.5	80	84.5
Pancreas	651	54	64	71	78	83
Larynx	111	52	58	64	73	82
Lung	1,909	55	62	68	75	80
Breast	2,549	43	51	61	73	81
Cervix	159	37	48	65	77	82
Corpus	576	53	59	67	76	81
Ovaries	436	45	55	64	75	81
Prostate	1,318	62	67	74	79	84
Testis	61	22	26	35	43	57
Kidneys	571	50	59	67	75	80
Urinary bladder	650	54	63	71	78	83.5
Melanoma	474	35	44	58	72	79
Brain	304	31	40	55	67	74
Thyroid gland	268	29	39	50.5	67	78
Connective tissue	121	32	43	59	77	84
Leukemias	350	41	56	67	77	83
Lymphomas	411	38	51	64	75	81
Hematopoietic system	242	55	63	69	77	82

well known for its data quality and completeness (5). Our empirical analysis was carried out among patients aged 15 or more years with a first diagnosis of cancer in 1989. We calculated 5- and 10-year absolute, expected, and relative survival rates and their standard errors for each of 25 common forms of cancer separately. Expected survival rates were calculated from age-, sex-, and calendar period-specific population life tables of Finland (which were assumed to have negligible random error) using Hakulinen's method (6).

The standard errors of survival rates were determined in two ways. First, the standard error of the absolute survival rates was calculated according to Greenwood's method (7), the standard errors of the expected survival rates were assumed to be zero, and the standard errors of the relative survival rates were obtained by dividing the standard errors of the absolute survival rates by the expected survival rates. This approach reflects common practice in the analysis of population-based cancer registry data. Second, all standard errors were estimated nonparametrically by bootstrap analysis as the standard deviation of the respective point estimates obtained in 10,000 bootstrap samples (full resamples) for each cancer site (8). The standard errors obtained by the bootstrap analysis and those obtained by the standard procedure were compared by calculating their ratios. Furthermore, Pearson's correlation coefficients of the point estimates of absolute and expected survival rates in the bootstrap samples were calculated to assess the correlation of sampling errors in the absolute and expected survival rates.

All survival analyses were carried out with publicly available SAS (SAS Institute, Inc., Cary, North Carolina) macros for both absolute and relative survival rates (9, 10), which were extended to allow for bootstrap analyses (a copy of the extended program may be obtained from the first author by request).

RESULTS

Table 1 shows the numbers and age distribution of cancer patients included in the analysis by cancer site. Breast cancer was the most common form of incident cancer in Finland in 1989, followed by lung cancer, prostate cancer, stomach cancer, and colon cancer. The median age at diagnosis was above 60 years for most of the cancers

included in the analysis, and it ranged from 74 years for patients with prostate cancer to 35 years for patients with cancer of the testis. Furthermore, there was a wide age range of patients even within cancer sites.

The point estimates and their standard errors of 5- and 10-year absolute survival rates are shown in table 2. Prognosis strongly varied by cancer site, with 5-year and 10-year absolute survival rates ranging from 88.5 percent and 86.9 percent for patients with testicular cancer to 1.4 percent and 1.1 percent for patients with pancreatic cancer. Standard errors for 5-year and 10-year absolute survival rates also strongly vary by cancer site. The standard errors determined by Greenwood's method and by bootstrap analysis are generally very similar (within ± 1 percent in all cases for 5-year absolute survival rates and in 21 of 25 cases for 10-year absolute survival rates).

Table 2 also shows the point estimates and their standard errors according to bootstrap analysis of 5-year and 10-year expected survival rates. The point estimates are highest for patients with testicular cancer because of their relatively young age and lowest for patients with prostate cancer because of their relatively old age. Although the bootstrap estimates of the standard errors of expected survival are lower than those of absolute survival for most forms of cancer, they are not entirely negligible and range from 0.31 to 2.24 percent and from 0.52 to 3.48 percent for 5-year and 10-year expected survival, respectively. Table 2 also shows that the correlation between absolute and expected survival rates in the bootstrap samples is substantial. Pearson's correlation coefficients varied from 0.11 to 0.76 for 5-year survival and from 0.09 to 0.76 for 10-year survival. The highest correlations are seen for cancers of the testis, thyroid gland, cervix and corpus uteri, and urinary bladder and for melanoma.

As table 3 shows, differences between standard errors for 5-year and 10-year relative survival rates derived by the standard procedure and by bootstrap analysis are often much larger than the very small differences seen for 5-year and 10-year absolute survival rates. With few exceptions, the standard errors obtained by the standard procedure exceed the standard errors obtained by bootstrap analysis. The former are up to 17 percent higher than the latter for 5-year relative survival, and the differences range up to 32 percent for 10-year relative survival rates. For nine of 25 cancers, the differences in standard errors of 5-year relative survival are 5 percent or higher. For three cancers, the differences are 10 percent or higher. Pertinent differences in standard errors of 10-year relative survival rates are seen for 15 and seven forms of cancer, respectively. The differences between the two types of standard errors are largest for those cancers with the highest correlation between absolute and expected survival rates, most notably for cancers of the thyroid gland and the testis.

To assess reproducibility of the observed patterns in different samples, we repeated all analyses separately for patients diagnosed in 1988 and 1987. Despite some variation in the survival rates between the years, the patterns regarding the standard errors were generally very similar, and results are therefore not shown separately to save space.

DISCUSSION

This analysis demonstrates that the standard errors of relative survival rates may be overestimated to some non-negligible extent with traditional methods of analysis. Potential overestimation appears to be more pronounced for 10-year relative survival rates than for 5-year relative survival rates, and it ranged up to 32 percent in our empirical evaluation for patients diagnosed with 25 common forms of cancer in Finland in 1989.

The rationale behind traditional estimation of the standard error of relative survival rates is that random variation in expected survival rates is negligible for practical purposes, as the latter are determined from population mortality figures that are typically based on very large numbers. However, even though random variation of the latter can typically be neglected (an assumption also made in our bootstrap analyses), the expected survival of a sample of cancer patients is still dependent on the age distribution of this sample, which is subject to random variation. As a result, there is also random variation in the expected survival of samples of cancer patients.

In theory, the additional random variation in expected survival rates might increase or reduce the standard error of relative survival rates (or leave them unchanged). More formally, let A and E denote absolute and expected survival with variances $\mathrm{var}(A)$ and $\mathrm{var}(E)$, respectively, and covariance $\mathrm{cov}(A,E)$. Then, applying the delta method, the variance of relative survival, $\mathrm{var}(R)$, is given as

$$\mathrm{var}(R) = (A/E)^2 \times (\mathrm{var}(A)/A^2 + \mathrm{var}(E)/E^2 - 2 \times \mathrm{cov}(A,E)/(A \times E)).$$

If random variation in absolute and expected survival rates were independent or negatively correlated, one might expect random error of relative survival rates to be higher than in the absence of random error in expected survival rates. However, a positive correlation in random variation of absolute and expected survival rates was observed for all forms of cancer assessed in this analysis. This result is not surprising, because both absolute and expected survival rates decrease with increasing age, and random variation in the age distribution of the samples of cancer patients would typically alter absolute and expected survival rates in the same direction. According to our analyses, this positive correlation is strong enough that relative survival rates typically have lower standard errors than suggested by conventional modes of calculation, which ignore random variation in expected survival rates altogether.

We are not aware of a possibility to calculate standard errors of relative survival rates that takes the random variation of both absolute and expected survival rates and their correlation into account, in a straightforward manner. Such standard errors can be empirically derived, however, by resampling techniques, of which bootstrap analysis is a particularly useful method (8). Application of resampling techniques in this context requires that the contributions of all patients to the calculations of expected survival are determined on an individual basis, an approach used in our recently developed computer programs for relative survival

TABLE 2. Absolute and expected survival after 5 years and 10 years of patients aged 15 or more years with a first diagnosis of common forms of cancer, Finland, 1989

	Absolute survival (%)								Expected survival (%)				Correlation*	
	5 years				10 years				5 years		10 years		5 years	10 years
	PE†	SE_G†	SE_B†	SE_G/SE_B ratio‡	PE	SE_G	SE_B	SE_G/SE_B ratio	PE	SE_B	PE	SE_B		
Oral cavity	48.52	2.45	2.45	1.00	33.31	2.31	2.32	1.00	78.79	0.93	61.10	1.39	0.26	0.41
Esophagus	5.50	1.61	1.60	1.01	3.50	1.30	1.29	1.01	71.89	1.44	50.06	1.95	0.11	0.09
Stomach	17.09	1.17	1.17	1.00	13.04	1.05	1.04	1.00	75.08	0.63	55.03	0.90	0.22	0.24
Small intestine	36.62	5.72	5.66	1.01	22.54	4.96	4.95	1.00	81.32	2.24	65.62	3.48	0.24	0.37
Colon	40.21	1.57	1.59	0.99	29.89	1.47	1.47	1.00	76.90	0.61	57.38	0.91	0.30	0.36
Rectum	34.89	1.95	1.95	1.00	24.71	1.76	1.75	1.00	77.54	0.75	58.21	1.13	0.23	0.33
Liver	6.28	1.17	1.17	1.00	3.95	0.94	0.96	0.98	75.05	0.93	54.30	1.36	0.16	0.14
Pancreas	1.38	0.46	0.46	0.99	1.08	0.40	0.40	1.00	76.78	0.73	56.99	1.07	0.12	0.13
Larynx	46.85	4.74	4.71	1.01	31.53	4.41	4.38	1.01	79.64	1.75	62.67	2.57	0.35	0.37
Lung	7.91	0.62	0.61	1.01	4.45	0.47	0.46	1.02	78.97	0.35	59.71	0.53	0.15	0.17
Breast	71.19	0.90	0.90	0.99	54.18	0.99	0.99	1.00	88.13	0.31	76.57	0.52	0.33	0.42
Cervix	49.06	3.96	3.98	1.00	42.14	3.92	3.94	1.00	84.98	1.41	71.27	2.34	0.52	0.57
Corpus	62.85	2.01	2.00	1.01	51.36	2.08	2.10	0.99	85.27	0.65	70.38	1.07	0.42	0.48
Ovaries	32.34	2.22	2.24	0.99	24.31	2.05	2.05	1.00	86.64	0.77	73.74	1.26	0.26	0.31
Prostate	42.60	1.36	1.37	0.99	19.14	1.08	1.08	1.00	67.70	0.49	43.02	0.65	0.21	0.27
Testis	88.52	4.08	4.10	0.99	86.85	4.33	4.36	0.99	96.49	0.88	92.62	1.70	0.76	0.71
Kidneys	44.83	2.08	2.08	1.00	32.57	1.96	1.93	1.02	82.18	0.65	65.13	1.05	0.27	0.34
Urinary bladder	49.23	1.96	1.99	0.99	31.63	1.83	1.88	0.97	74.41	0.74	53.60	1.09	0.39	0.51
Melanoma	71.30	2.08	2.05	1.01	59.03	2.26	2.25	1.01	87.58	0.70	75.88	1.17	0.41	0.48
Brain	23.03	2.41	2.40	1.01	15.46	2.07	2.06	1.01	90.80	0.71	81.32	1.23	0.30	0.30
Thyroid gland	77.22	2.56	2.57	1.00	72.71	2.72	2.74	1.00	91.34	0.83	82.72	1.46	0.63	0.70
Connective tissue	43.80	4.51	4.49	1.01	31.40	4.22	4.19	1.01	83.03	1.89	70.23	2.86	0.13	0.24
Leukemias	31.43	2.48	2.47	1.01	17.71	2.04	2.06	0.99	80.35	1.02	63.81	1.54	0.31	0.34
Lymphomas	37.23	2.38	2.39	1.00	23.84	2.10	2.07	1.01	83.98	0.85	69.38	1.35	0.38	0.35
Hematopoietic system	20.25	2.58	2.57	1.00	5.37	1.45	1.44	1.01	78.79	1.12	59.79	1.72	0.31	0.24

* The correlation of absolute and expected survival rates in 10,000 bootstrap samples.

† PE, point estimate; SE_G, standard error obtained according to Greenwood's method (absolute survival only); SE_B, standard error obtained by bootstrap analysis (absolute and expected survival).

‡ Absolute survival only.

TABLE 3. Relative survival after 5 years and 10 years of patients aged 15 or more years with a first diagnosis of common forms of cancer, Finland, 1989

	Relative survival (%)							
	5 years				10 years			
	PE*	SE_S*	SE_B*	SE_S/SE_B ratio	PE	SE_S	SE_B	SE_S/SE_B ratio
Oral cavity	61.58	3.11	3.01	1.03	54.51	3.79	3.48	1.09
Esophagus	7.65	2.24	2.22	1.01	6.99	2.60	2.57	1.01
Stomach	22.76	1.56	1.53	1.02	23.69	1.90	1.84	1.03
Small intestine	45.05	7.03	6.77	1.04	34.35	7.56	7.07	1.07
Colon	52.28	2.05	1.98	1.03	52.09	2.56	2.39	1.07
Rectum	45.00	2.51	2.46	1.02	42.45	3.03	2.85	1.06
Liver	8.37	1.56	1.55	1.01	7.28	1.73	1.75	0.99
Pancreas	1.80	0.60	0.60	0.99	1.89	0.71	0.71	1.00
Larynx	58.80	5.95	5.60	1.06	50.28	7.03	6.53	1.08
Lung	10.02	0.78	0.77	1.01	7.46	0.79	0.77	1.03
Breast	80.78	1.02	0.97	1.05	70.75	1.29	1.18	1.10
Cervix	57.72	4.66	4.27	1.09	59.11	5.49	4.69	1.17
Corpus	73.71	2.36	2.17	1.09	72.99	2.96	2.64	1.12
Ovaries	37.32	2.59	2.51	1.03	32.97	2.79	2.66	1.05
Prostate	62.92	2.01	1.98	1.02	44.47	2.52	2.42	1.04
Testis	91.74	4.23	3.66	1.15	93.77	4.68	3.70	1.26
Kidneys	54.56	2.53	2.44	1.04	50.01	3.01	2.79	1.08
Urinary bladder	66.16	2.64	2.49	1.06	59.02	3.41	3.04	1.12
Melanoma	81.41	2.37	2.16	1.10	77.79	2.98	2.61	1.14
Brain	25.36	2.66	2.59	1.03	19.01	2.55	2.46	1.04
Thyroid gland	84.53	2.81	2.41	1.17	87.89	3.29	2.49	1.32
Connective tissue	52.75	5.43	5.38	1.01	44.70	6.01	5.81	1.03
Leukemias	39.11	3.09	2.96	1.04	27.75	3.20	3.07	1.04
Lymphomas	44.33	2.84	2.71	1.05	34.37	3.03	2.83	1.07
Hematopoietic system	25.70	3.28	3.17	1.03	8.98	2.42	2.36	1.03

* PE, point estimate; SE_S, standard error according to standard procedure; SE_B, standard error according to bootstrap analysis.

(9, 10), and not on an aggregate basis, as it had mostly been done in previous commonly used computer programs, such as that by Hakulinen and Abeywickrama (11).

A disadvantage of resampling techniques is that a very large number of replications is necessary to come up with reliable estimates of standard errors, requiring substantial computing time. In our analyses, we decided to use 10,000 replications per analysis to ensure very high levels of reliability. The almost perfect agreement of standard errors obtained by Greenwood's formula and by bootstrap analysis for the absolute survival rates in all cases reassures us that this number of replications is sufficient to come up with reliable bootstrap estimates. It took about 4 hours to run all the bootstrap analyses shown in this paper on a standard laptop (Pentium 4 processor (Intel Corporation, Santa Clara, California), 3.06 GHz, 512 MB RAM). Obviously, there is a tradeoff between computation time and precision of the estimates, and lower numbers of bootstrap replications may often be sufficient for practical purposes. When the analyses were repeated with 1,000 bootstrap replications per analysis, the agreement of standard errors obtained by Greenwood's formula and by bootstrap analysis for the absolute survival rates was still within ± 5 percent in all cases (and within ± 2 percent in 40 of 50 cases), and the computing time was reduced to about half an hour. Although these computing times appear to be quite long, computing time most likely will further rapidly decrease with more powerful hardware equipment in the future, and the relevance of computing time (which can usually be easily allocated during nights or weekends) appears to be very small if compared with the time and effort it takes to collect pertinent cancer registry data.

In summary, our analyses suggest that conventional estimates of standard errors for relative survival rates may often be too high, particularly for long-term relative survival rates. Although more time and computer intensive, bootstrap analysis may be a useful tool to estimate standard errors of relative survival rates.

REFERENCES

1. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. Natl Cancer Inst Monogr 1961;6:101–21.
2. Parkin DM, Hakulinen T. Cancer registration: principles and methods. Analysis of survival. IARC Sci Publ 1991;(95): 159–76.
3. Estève J, Benhamou E, Raymond L. Statistical methods in cancer research. Vol IV. Descriptive epidemiology. IARC Sci Publ 1994;(128):1–302.
4. Mathew A, Pandey M. Attributing death to cancer: cause-specific survival estimation. J Postgrad Med 2002;48: 322–6.
5. 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.
6. Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. Biometrics 1982;39:933–42.
7. Greenwood M. The natural duration of cancer. In: Reports on public health and medical subjects. London, England: His Majesty's Stationery Office,1926:1–26.
8. Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Stat Sci 1986;1:54–77.
9. Brenner H, Hakulinen T, Gefeller O. Computational realization of period analysis for monitoring cancer patient survival. Epidemiology 2002;13:611–12.
10. Arndt V, Talbäck M, Gefeller O, et al. Modification of SAS macros for more efficient analysis of relative survival rates. Eur J Cancer 2004;40:778–9.
11. Hakulinen T, Abeywickrama KH. A computer program package for relative survival analysis. Comput Programs Biomed 1985;19:197–207.