ELSEVIER

European Journal of Cancer 40 (2004) 2317–2322

European Journal of Cancer

www.ejconline.com

An alternative approach to age adjustment of cancer survival rates

Hermann Brenner [a,*], Volker Arndt [a], Olaf Gefeller [b], Timo Hakulinen [c,d]

[a] *Department of Epidemiology, German Centre for Research on Ageing, Bergheimer Strasse 20, D-69115 Heidelberg, Germany*
[b] *Department of Medical Informatics, Biometry and Epidemiology, University of Erlangen-Nuremberg, Waldstrasse 6, D-91504 Erlangen, Germany*
[c] *Finnish Cancer Registry, Liisankatu 21 B, FIN-00170 Helsinki, Finland*
[d] *Department of Public Health, University of Helsinki, FIN-00170 Helsinki, Finland*

Received 30 March 2004; received in revised form 21 June 2004; accepted 13 July 2004
Available online 18 August 2004

Abstract

The measures of cancer prognosis most commonly reported by cancer registries are 5- or 10-year absolute or relative survival rates. Because cancer survival rates often vary by the age of cancer patients, and because the age structure of cancer patients often varies between populations, age adjustment is crucial for comparative analyses of cancer survival rates. However, traditional age adjustment often breaks down for long-term survival rates, particularly for sparse data, and it may provide inconsistent results for relative survival rates, even if age adjustment is made to the study population's own age distribution. In this manuscript, we propose an alternative approach to age adjustment of both absolute and relative survival rates to overcome both the practical and conceptual problems inherent in traditional age adjustment. We outline the computational realisation of this approach, and we give an empirical illustration of its application using data from the nationwide Finnish Cancer Registry.

Keywords: Age adjustment; Biostatistics; Cancer registries; Epidemiological methods; Survival; Time trends

1. Introduction

The measures of cancer prognosis most commonly reported by cancer registries are long-term survival rates, such as 5- or 10-year survival rates [1]. Typically, both absolute (observed) and relative survival rates are reported. Relative survival rates are derived as the ratios of the observed survival rates and the expected survival rates in the absence of cancer in a population of comparable age and gender distribution (the latter are usually estimated from population life tables). The relative survival rates can be thought of as "net" measures of cancer patient survival indicating cancer patient survival in the hypothetical situation in which cancer is the only cause of death [2].

Absolute, and, to a lesser extent, relative survival rates often vary by the age of cancer patients. Therefore, comparisons of cancer patient survival between populations or at various calendar periods within one population may be confounded by age, if there are differences in the age distribution of cancer patients. To overcome this problem, age standardisation is commonly employed in comparative analyses of cancer survival rates, such as the EUROCARE project [3–6]. Traditionally, age-standardised survival rates have been calculated as a weighted average of age-specific survival rates within subgroups of patients defined by age at diagnosis, with weights equal to the proportion of patients in those subgroups in some standard population, such as the European standard cancer population or the world standard cancer population [3–7]. A major problem in this context is often the sparseness of data within certain (mostly older) age groups, which may hinder the calculation of age-specific survival rates. Furthermore, it has

* Corresponding author. Tel.: +49-6221-548140; fax: +49-6221-548142.
E-mail address: brenner@dzfa.uni-heidelberg.de (H. Brenner).

0959-8049/$ - see front matter
doi:10.1016/j.ejca.2004.07.007

been shown that, for relative survival rates, this procedure may provide adjusted survival rates which are conceptually inconsistent with the crude rates and may substantially differ from the latter, even if adjustment is made to the study population's own age distribution [8].

In this manuscript, we propose an alternative approach to age adjustment of both absolute and relative survival rates to overcome both the practical and the conceptual problems inherent in traditional age adjustment. We outline a computational realisation of this approach, which does not require to carry out individual age-specific survival analyses, and we give an empirical illustration of its application using data from the nationwide Finnish Cancer Registry.

2. Our alternative approach

When employing traditional age adjustment one first calculates the age-specific estimates of survival in different age groups, and then combines these estimates in a weighted average (with weights reflecting the age distribution of the standard population). In our alternative approach, specific weights are first individually assigned to all patients in different age groups, and one then carries out conventional survival analysis using the "weighted individual data", in which the weights are applied to the contributions of patients to the numbers of persons at risk and death. Whereas in the unadjusted (crude) analysis each patient in the study population and her/his contributions to the numbers of persons at risk and deaths are (implicitly) entered with a weight of 1, the proposed form of age adjustment gives weights higher (lower) than 1 to patients in age groups which are underrepresented (overrepresented) in the study population compared with the standard population.

More formally, let r_i be the relative proportion of patients in age group i in the study population of total sample size n, and s_i the corresponding proportion in the standard population. Then, each patient (whether alive during the full follow-up period of interest, censored or dead) and all of her/his contributions to the numbers of persons at risk and deaths are assigned a weight of s_i/r_i. This procedure implies, firstly, that the weighted study population has exactly the same age distribution as the standard population (as the sum of weights in each age group i equals $nr_i(s_i/r_i) = ns_i$), and, secondly, that the total sum of weights equals the total number n of study participants (as in the crude analysis), as

$$\sum_i ns_i = n.$$

Compared with traditional age adjustment of survival rates, this procedure has two obvious advantages: firstly, and probably of most practical importance, it does not break down if none of the patients within one or more age groups is followed over the entire follow-up period of interest, a problem commonly encountered in traditional age adjustment. Secondly, this procedure also assures, that age adjustment of relative survival rates to the study population's own age distribution, which assigns a weight of 1 to each patient, yields exactly the same result as obtained in the crude analysis. This 'natural' property is usually not fulfilled for traditionally adjusted relative survival rates as previously shown in [8].

3. Computational realisation

For easy implementation of the alternative method proposed in this paper to derive adjusted survival rates, we have extended recently developed SAS macros for both relative and absolute survival rates [9,10]. These macros, as well as their extension described below, may not only be used for traditional "cohortwise" analysis of survival rates, but also for "period analysis", a more recently introduced method to derive more up-to-date long-term survival estimates [11].

A detailed description of the previously available macros, which can be downloaded free of charge from the Internet, has been given elsewhere in Refs. [9–11]. Briefly, two macros have been developed, one called "period" in which relative survival rates are calculated according to the Ederer II method [12], and one called "periodh" in which they are calculated according to Hakulinen's method [13]. We now provide extensions of these macros, denoted "adperiod" and "adperiodh", in which all contributions of study participants to the observed and expected person–time at risk and to the numbers of deaths are weighted. Thus, the only additional step to be done by the investigator is *a priori* assignment of appropriate weights to the study participants as described above. The macros will then provide estimates of the adjusted survival rates in a one-step analysis. In particular, there is no need to carry out individual age specific-survival analyses.

To facilitate application, pertinent SAS programs can be found on our web-site http://www.imbe.med.uni-erlangen.de/issan/SAS/period/period.htm. The site also provides the necessary preparatory steps and an exemplary SAS program for specification of the age structure of the standard population, a macro denoted "adweight" that allows for convenient, flexible assignment of appropriate weights to study participants, as well as the extended, commented macros for carrying out weighted survival analyses are also listed. All programs can be downloaded free of charge from the statistical archive network maintained by the Department of Medical Informatics, Biometry and Epidemiology at the University of Erlangen-Nuremberg (http://www.imbe.med.uni-erlangen.de/issan/SAS/period/period.htm) or obtained on diskette by written request to the first author. The source code of the SAS macros is open code under the conditions of the GNU-General Public Licence [14] and thus can be modified by future users to adapt the macros to specific needs.

4. Empirical illustration

The alternative method of age adjustment proposed in this paper is illustrated for an analysis of relative survival of patients with cancer of the corpus uteri in Finland. The database for this analysis is the nationwide Finnish Cancer Registry, which is well known for its high quality and completeness of data and its long time period of cancer registration of more than 50 years [3,15]. The aim of this analysis was to compare 10-year relative survival of patients diagnosed in 1985–1989 (the most recent cohort of patients for whom 10-year follow-up was complete at the time of this analysis) with the 10-year relative survival of patients diagnosed 30 years earlier, i.e., in 1955–1959. Only patients with a first diagnosis of cancer of the corpus uteri above age 15 years were included, and we excluded patients for whom the death certificate or an autopsy report was the only notification to the cancer registry (0.9% and 0.7% of patients, respectively). Relative survival rates were calculated according to Hakulinen's method [13].

The numbers and age distribution of patients diagnosed in both time periods are shown in Table 1. Among the 1071 patients diagnosed in 1955–1959, only 23.9% and 6.3% were 65 years or older, and 75 years or older, respectively. These proportions increased to 55.2% and 24.2%, respectively, among the 2511 patients diagnosed in 1985–1989. Furthermore, 10-year relative survival strongly decreased with age for patients diagnosed in both time periods. Despite the much higher average age of patients diagnosed in 1985–1989 compared with patients diagnosed in 1955–1959, even overall (crude) 10-year relative survival increased from 63.9% for the former group to 75.1% for the latter group. However, within each single age group a more pronounced increase in relative survival was observed. Taken together, these patterns point to the need of age adjustment of single summary measures of 10-year relative survival to disclose the full extent of improvements in relative survival over time.

In the following, various options of age adjustment are outlined using stratification by the five age groups shown in Table 1. An obvious way of age adjustment in a comparison of 10-year relative survival of the two patient groups would be to adjust 10-year relative survival of one of the patients groups to the age distribution of the other group.

As Table 2 shows, the 10-year relative survival estimate of patients diagnosed in 1985–1989 would become substantially higher than the estimate obtained in the crude analysis if it was adjusted to the age distribution of patients diagnosed in 1955–1959 in the traditional manner, i.e., by calculating a weighted average of the age-specific survival rates, with weights equal to the proportions of patients in the various age groups in 1955–1959. However, comparison of this adjusted 10-year survival estimate to the crude survival estimate of patients diagnosed in 1955–1959 (difference: 15.1% units) would still not disclose the full extent of the improvement, as the relative survival estimate of the latter patients would also be altered (here: reduced) if it was adjusted to this group's own age distribution in the traditional manner. Although both adjusted survival estimates may be used to come up with a valid comparison, none of these estimates actually coincides with the crude estimate for either group of patients.

Conversely, the 10-year relative survival estimate for patients diagnosed in 1955–1959 would become substantially lower than the crude relative survival estimate as expected if it was adjusted to the age distribution of patients diagnosed in 1985–1989 in the traditional manner, suggesting a tremendous increase in 10-year relative survival of 26.5% units if a comparison was made to the crude relative survival estimate for patients diagnosed in 1985–1989. However, in this case, the true extent of improvement would be overestimated, as the 10-year relative survival estimate for the latter group would likewise become lower, albeit to a much lesser extent, if it was adjusted to this population's own age structure.

Table 1
Age distribution and 10-year relative survival rates (RSR) by age of patients diagnosed with endometrial cancer in Finland in 1955–1959 and 1985–1989

Age at diagnosis (years)	Patients diagnosed in 1955–1959			Patients diagnosed in 1985–1989		
	N	%	10-Year RSR (%)	*N*	%	10-Year RSR (%)
15–44	70	6.5	78.0	84	3.3	93.3
45–54	324	30.3	72.8	342	13.6	87.3
55–64	421	39.3	63.1	700	27.9	77.6
65–74	188	17.6	38.1	778	31.0	74.0
75+	68	6.3	27.6	607	24.2	47.3
Total	1071	100	63.9	2511	100	75.1

Table 2
Comparison of 10-year relative survival of cohorts of patients diagnosed with endometrial cancer in Finland in 1955–1959 and 1985–1989 using crude analysis and traditional age adjustment

Type of analysis	10-Year relative survival (%)		
	1955–1959 Cohort	1985–1989 Cohort	Difference
No age adjustment (crude survival)	63.9	75.1	+11.2
Traditional age adjustment			
1985–1989 Cohort adjusted to 1955–1959 age structure	63.9	79.0	+15.1
Both cohorts adjusted to 1955–1959 age structure	60.4	79.0	+18.6
1955–1959 Cohort adjusted to 1985–1989 age structure	48.6	75.1	+26.5
Both cohorts adjusted to 1985–1989 age structure	48.6	71.0	+22.4

Table 3
Weighting of patients diagnosed with endometrial cancer in Finland 1955–1959 and 1985–1989 in crude analysis and with alternative age adjustment

Age at diagnosis (years)	Patients diagnosed in 1955–1959			Patients diagnosed in 1985–1989		
	Crude analysis	Adjustment to 1955-1959 age structure	Adjustment to 1985–1989 age structure	Crude analysis	Adjustment to 1955–1959 age structure	Adjustment to 1985–1989 age structure
15–44	1	6.5/6.5 = 1	3.3/6.5 = 0.51	1	6.5/3.3 = 1.95	3.3/3.3 = 1
45–54	1	30.3/30.3 = 1	13.6/30.3 = 0.45	1	30.3/13.6 = 2.22	13.6/13.6 = 1
55–64	1	39.3/39.3 = 1	27.9/39.3 = 0.71	1	39.3/27.9 = 1.41	27.9/27.9 = 1
65–74	1	17.6/17.6 = 1	31.0/17.6 = 1.77	1	17.6/31.0 = 0.57	31.0/31.0 = 1
75+	1	6.3/6.3 = 1	24.2/6.3 = 3.81	1	6.3/24.2 = 0.26	24.2/24.2 = 1
Sum of weights	1071	1071	1071	2511	2511	2511

Table 4
Comparison of 10-Year relative survival of cohorts of patients diagnosed with endometrial cancer in Finland in 1955–1959 and 1985–1989 using crude analysis and alternative age adjustment

Type of analysis	10-Year relative survival (%)		
	1955–1959 Cohort	1985–1989 Cohort	Difference
No age adjustment (crude survival)	63.9	75.1	+11.2
Alternative age adjustment:			
to 1955–1959 age structure	63.9	80.6	+16.7
to 1985–1989 age structure	55.6	75.1	+19.5
to EUROCARE-2 age structure	57.5	76.3	+18.8

By contrast, application of the alternative age adjustment procedure outlined in this paper, which can be easily implemented by assigning weights to the patients within different age groups as outlined in Table 3 and by using the extended macro *adperiodh* given on our web-site, does not alter the crude relative survival rates if adjustment is made to a study population's own age structure. Hence, the relative survival estimate obtained for the 1985–1989 cohort by adjustment to the age structure of the 1955–1959 cohort can be readily compared with the crude relative survival rate of the latter and interpreted as the relative survival rate that would have been expected for the 1985–1989 cohort if this cohort had had the same age distribution as the 1955–1959 cohort (see Table 4). Analogously, the relative survival estimate obtained for the 1955–1959 cohort by adjustment to the age structure of the 1985–1989 cohort can be readily compared with the crude relative survival estimate of the latter and interpreted as the relative survival rate that would have been expected for the 1955–1959 cohort if this cohort had had the same age distribution as the 1985–1989 cohort.

Obviously, the alternative age adjustment can also be applied with some external standard. For illustration, Table 4 therefore also includes 10-year relative survival rates for both cohorts of patients after adjustment to the age structure of the EUROCARE-2 standard population (reflecting the age structure of patients diagnosed with endometrial cancer in 1985–1989 who were included in the EUROCARE project) [4]. Overall, results were quite similar to those of age adjustment to the 1985–1989 age structure of the Finnish women.

5. Discussion

In this paper, we propose a simple, alternative method for age adjustment of cancer survival. We outline its computational realisation and we illustrate its use with an application to data from the nationwide Finnish Cancer Registry.

Our alternative method overcomes essential shortcomings of the traditional method for age adjustment of relative survival rates. The latter often breaks down if data within single age groups are sparse, in which case it may not be possible to derive age-specific estimates of long-term cancer survival. This problem is often encountered in older age groups, particularly if long-term survival rates are calculated, and it is due to the rapid decline and eventual disappearance of the numbers of persons at risk during follow-up. By contrast, our alternative method of age adjustment is unaffected by this problem as there is no need to calculate age-specific survival rates. The only precondition here is that there is at least one patient in each age group at the beginning of follow-up (i.e., at the time of diagnosis), which is the minimum requirement for calculating the weights to be assigned to each patient, and which is a much less stringent precondition than the presence of at least one patient at risk at each time interval of follow-up needed in traditional age adjustment (unless all patients have died before).

Another major shortcoming of traditional age adjustment is that the adjusted relative survival rates are conceptually different from the crude relative survival rates, and that both measures may differ substantially, even if adjustment is made to the study population's own age structure. As a result, the common practice of making survival estimates between two cancer populations comparable by adjusting the survival estimate of one of the populations to the age distribution of the other one (used as standard) may often be misleading unless the survival estimate in the standard population is also adjusted to its own age distribution. The practical relevance of this problem, which becomes increasingly important with increasing length of follow-up, has been previously demonstrated for a large number of different cancers [8]. It was also very evident in the empirical example given in this paper. We have previously demonstrated a possible solution to overcome this problem by the use of alternative weights in traditional age adjustment of relative survival rates [8]. However, application of this solution was less straightforward, as it still required to carry out age-specific survival analyses, and it also required derivation of expected survival rates for the standard cancer population. The latter depend on the length of follow-up and may be difficult to obtain in some instances, e.g. when adjustment is made to some external standard, such as the world standard cancer population. The alternative method of age adjustment introduced in this paper is not affected by these limitations.

The empirical illustration given in this manuscript was restricted to 10-year relative survival rates calculated according to the Hakulinen method. However, the proposed alternative method is equally applicable to relative survival rates calculated by other methods, to calculation of absolute survival rates, or for different lengths of follow-up.

While the proposed alternative adjustment procedure has clear advantages in the comparison of survival between populations or between various calendar periods within one population, preferences may be different in situations where the focus is on comparisons of patient survival in various follow-up periods since diagnosis. As shown by Hakulinen more than 25 years ago [16], the age distribution of surviving cancer patients tends to change during follow-up (typically towards younger ages), and this change may confound comparisons of patient survival over time. This type of confounding would be eliminated or reduced by traditional age adjustment, where the weights for the various age groups remain the same in all of the follow-up periods.

In summary, we have outlined an alternative method of age adjustment of cancer survival rates which should facilitate valid comparisons of cancer survival rates between different populations or over time within one population even when data are sparse.

Conflict of Interest Statement

None declared.

Acknowledgement

The work of Timo Hakulinen was partly funded by the MaDaMe project of the Academy of Finland.

References

1. Parkin DM, Hakulinen T. Analysis of survival. In: Jensen OM, Parkin DM, MacLennan R, *et al.*, editors. *Cancer registration. Principles and methods*. IARC Scientific Publications No. 95, Lyon, International Agency for Research on Cancer, 1991, 159–176.
2. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. *Monogr Natl Cancer Inst* 1961, **6**, 101–121.
3. Berrino F, Sant M, Verdecchia A, *et al.*, editors. *Survival of cancer patients in Europe: the EUROCARE Study*. IARC Scientific Publications No. 132, Lyon, International Agency for Research on Cancer, 1995.
4. Berrino F, Capocaccia R, Estève J, editors. *Survival of cancer patients in Europe: the EUROCARE-2 study*. IARC Scientific Publications No. 151, Lyon, International Agency for Research on Cancer, 1999.
5. Coebergh JWW, Sant M, Berrino F, Verdecchia A. Survival of adult cancer patients in Europe diagnosed from 1978–89: the EUROCARE II study. *Eur J Cancer* 1998, **34**, 2137–2278.

6. Berrino F, Capocaccia R, Coleman MP, *et al*., editors. Survival of cancer patients in Europe: the EUROCARE-3 study. *Annals of Oncology* 2003, **14**(Suppl 5), v1–v155.
7. Sankaranarayanan R, Black RJ, Parkin DM, editors. *Cancer survival in developing countries*. IARC Scientific Publications No. 145. Lyon, International Agency for Research on Cancer, 1998.
8. Brenner H, Hakulinen T. On crude and age adjusted relative survival rates. *J Clin Epidemiol* 2003, **56**, 1185–1191.
9. Brenner H, Gefeller O, Hakulinen T. A computer program for period analysis of survival. *Eur J Cancer* 2002, **38**, 690–695.
10. Brenner H, Hakulinen T, Gefeller O. Computational realization of period analysis for monitoring cancer patient survival. *Epidemiology* 2002, **13**, 611–612.
11. 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**, 326–335.
12. 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.
13. Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics* 1982, **39**, 933–942.
14. Free Software Foundation. The GNU General Public License. Available from: www.gnu.org/licenses/licenses.html. Accessed September 12, 2003.
15. Teppo L, Pukkala E, Lehtonen M. Data quality and quality control of a population-based cancer registry. Experience in Finland. *Acta Oncol* 1994, **33**, 365–369.
16. Hakulinen T. On long-term relative survival rates. *J Chron Dis* 1977, **30**, 431–443.