EUROPEAN JOURNAL OF CANCER 47 (2011) 2202–2210

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

# Choosing the relative survival method for cancer survival estimation

*Timo Hakulinen* [a,*], *Karri Seppä* [a,b], *Paul C. Lambert* [c,d]

[a] *Finnish Cancer Registry, Institute for Statistical and Epidemiological Cancer Research, Pieni Roobertinkatu 9, FI-00130 Helsinki, Finland*
[b] *Department of Mathematical Sciences, University of Oulu, Oulu, Finland*
[c] *Department of Health Sciences, Centre for Biostatistics and Genetic Epidemiology, University of Leicester, 2nd Floor, Adrian Building, University Road, Leicester LE1 7RH, UK*
[d] *Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden*

ARTICLE INFO

*Article history:*
Received 22 November 2010
Received in revised form 7 March 2011
Accepted 9 March 2011
Available online 4 May 2011

*Keywords:*
Epidemiologic methods
Models
Neoplasms
Prognosis
Survival
Age standardisation

ABSTRACT

*Background:* The methods on how to calculate cumulative relative survival have been ambiguous and have given differences in empirical results.
*Methods:* The gold standard for the cumulative relative survival ratio is the weighted average of age-specific cumulative relative survival ratios, with weights proportional to numbers of patients at diagnosis. Mathematics and representative empirical materials from the population-based Finnish Cancer Registry were studied for the different relative survival methods and compared with the gold standard.
*Results:* The theoretical and empirical results show a good agreement between the method suggested in 1959 by Ederer and Heise (the so-called Ederer II method) and the gold standard. This result is in part due the fact that as follow-up time increases the conditional (annual) relative survival ratios become increasingly more independent of age. Moreover, the dependence between the excess mortality due to cancer and the baseline general mortality does not introduce an important enough selection in practice to cause a notable bias.
*Conclusion:* The use of the method by Ederer and Heise, multiplication of the annual relative survival ratios, instead of direct standardisation, should be considered in future applications. This would be particularly important for the long-term follow-up when age-specific relative survival is not available in the oldest age categories.


## 1. Introduction

Relative survival has been used by the world's population-based cancer registries for 60 years to give estimates of patient survival as far as their cancer is concerned, in the absence of other causes of death.[1,2] No cause-of-death information is needed as the mortality from other causes of death (expected mortality) is estimated from general population life tables.

Two ways of estimating the cumulative relative survival ratios were proposed by Ederer et al.: multiplication of the annual conditional relative survival ratios[3] and dividing the cumulative observed survival proportion by the cumulative expected proportion.[4] The results given by these two methods, often called the Ederer II and I methods, respectively, differ markedly, particularly when the length of follow-up exceeds ten years.[5,6] It has been shown that the results given by the Ederer I method converge with prolonged follow-up

* *Corresponding author:* Tel.: +358 9 135 331; fax: +358 9 135 5378.
E-mail address: timo.hakulinen@cancer.fi (T. Hakulinen).
0959-8049/$ - see front matter 
doi:10.1016/j.ejca.2011.03.011

towards the cumulative relative survival ratios of the youngest patients.[5] The "Hakulinen" method[6] is optimal under informative censoring when the relative survival is constant across the ages. However, it has often been applied even when this constancy assumption does not hold, which is the case with most cancers. Without applying regression modelling, the gold standard for cumulative relative survival ratio is the directly age-standardised cumulative relative survival ratio with weights proportional to the numbers of patients at the beginning of follow-up.[7] However, the gold standard is often not readily or reliably available, since long-term relative survival is not necessarily estimable or reliably estimable for all age groups, particularly the oldest age group.[5]

This study reports that the method proposed by Ederer and Heise[3] (the Ederer II method) works well for cumulative relative survival ratios and gives foundation for that finding. As a consequence, the standard reporting of cumulative relative survival ratios should be revised.

## 2. Patients and methods

Patients diagnosed in Finland in 1970–79 by the population-based Finnish Cancer Registry and followed until the end of 2009 were included in the analysis with stratification by age (groups 0–44, 45–64, 65–74 and 75 years and more), site and gender. In order to study the effect of censoring, the end of 1988 was used as the alternative end of follow-up. Cumulative

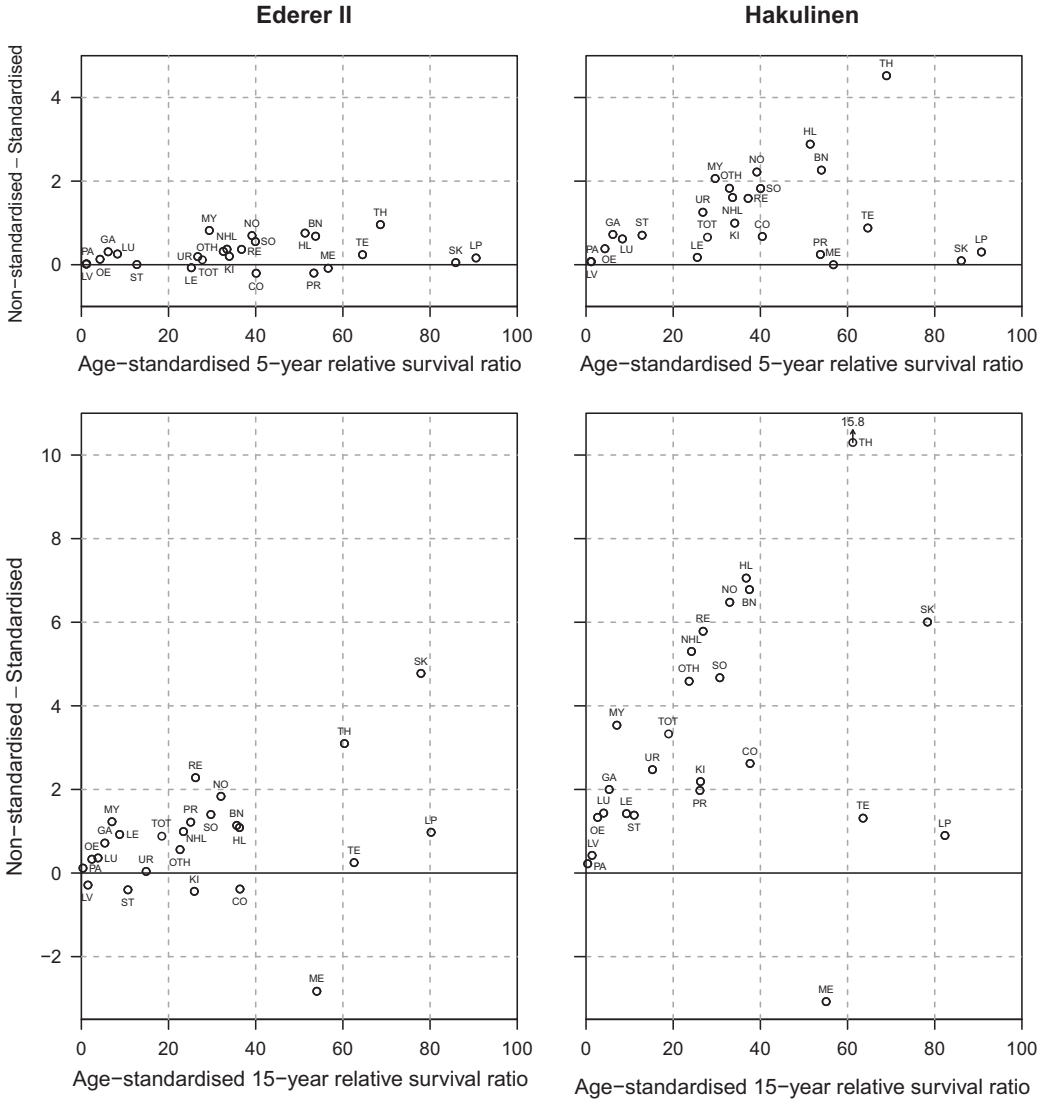


**Fig. 1 – Age-standardised five- and fifteen-year relative survival ratios (%) and differences between the non-standardised and age-standardised ratios (% units) obtained by the Ederer II (Ederer and Heise[3]) and Hakulinen methods (Hakulinen[6]) in males diagnosed with cancer in Finland in 1970–1979, by site. Results obtained by the Ederer I method (Ederer et al.[4]) equal to those given by the Hakulinen method. The site codes are: BN = brain, central nervous system; CO = colon; GA = gallbladder; HL = Hodgkin lymphoma; KI = kidney; LE = leukaemia; LP = lip; LU = lung; LV = liver; ME = skin, melanoma; MY = multiple myeloma; NHL = non-Hodgkin lymphoma; NO = nose; OE = oesophagus; OTH = other cancers; PA = pancreas; PR = prostate; RE = rectum; SK = skin, non-melanoma; SO = soft tissue; ST = stomach; TE = testis; TH = thyroid; TOT = all cancers combined; UR = urinary bladder.**

age-specific, overall non-standardised and age-standardised relative survival ratios were obtained by each of the methods proposed by Ederer et al.[3,4] and by the Hakulinen method.[6] Mathematical details were worked out for the methods in order to give background to the empirical findings.

## 3. Results

First, results with follow-up until the end of 2009 are considered. The patients thus have almost 30 years of practically complete follow-up with no censoring. Estimates from the Ederer I and Hakulinen methods are equal in the absence of censoring and so only the latter are reported.

For a five-year follow-up, it does not really matter how the relative survival is estimated (Figs. 1–3, Supplementary Table). When the follow-up gets longer, the age-standardised cumulative relative survival ratios based on all the methods agree fairly well, as opposed to the corresponding non-standardised ratios. A particularly good agreement prevails between the non-standardised and age-standardised ratios by using the Ederer II method. By considering the follow-up until the end of 1988 only, the resulting differences in censoring by age do not have an impact on these results (not shown).

The gold standard, being empirical, is also subject to random variation. Particularly the age-standardised relative survival ratio is prone to a large standard error due to small

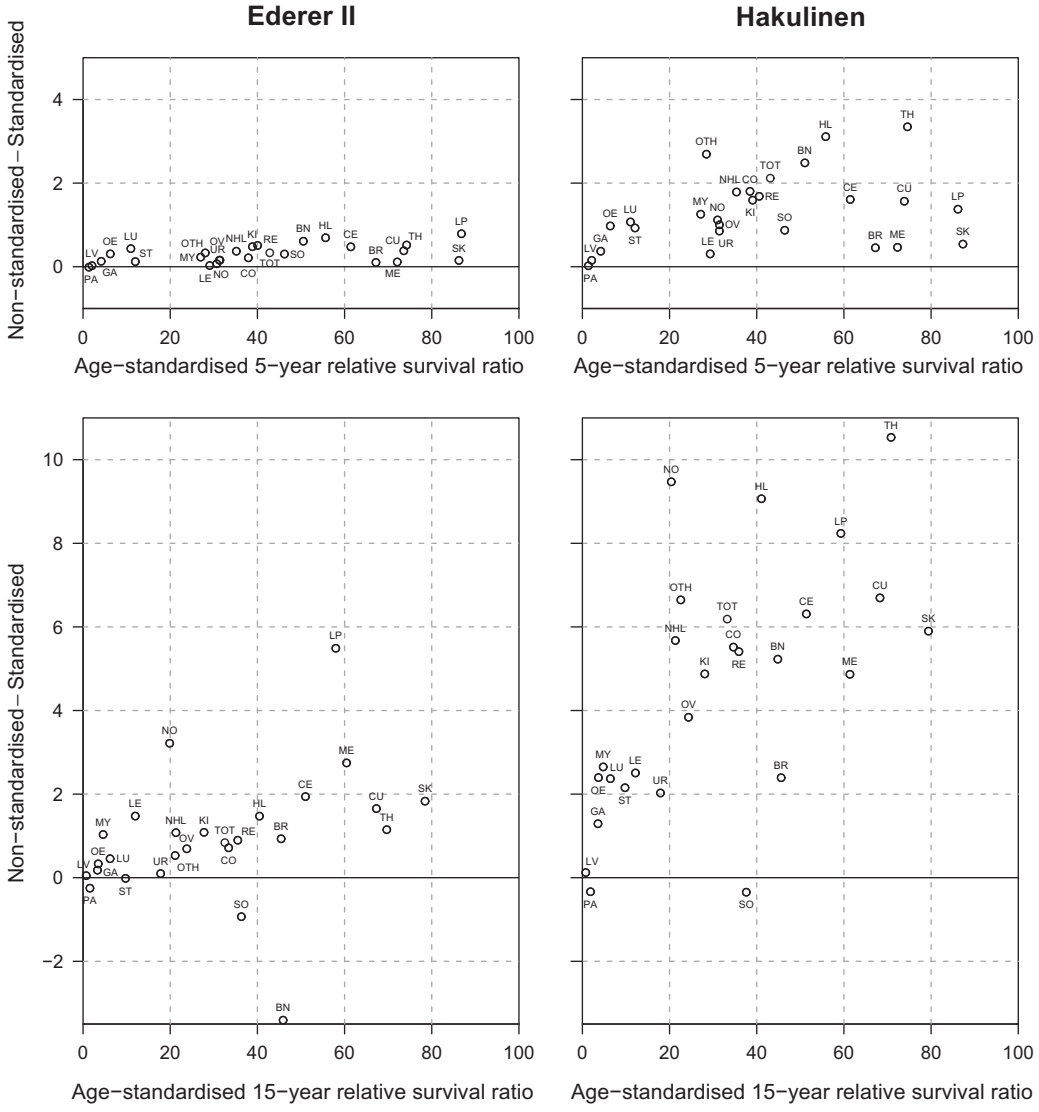


**Fig. 2 – Age-standardised five- and fifteen-year relative survival ratios (%) and differences between the non-standardised and age-standardised ratios (% units) obtained by the Ederer II (Ederer and Heise[3]) and Hakulinen methods (Hakulinen[6]) in females diagnosed with cancer in Finland in 1970–1979, by site. Results obtained by the Ederer I method (Ederer et al.[4]) equal to those given by the Hakulinen method. The site codes are: BN = brain, central nervous system; BR = breast; CE = cervix uteri; CO = colon; CU = corpus uteri; GA = gallbladder; HL = Hodgkin lymphoma; KI = kidney; LE = leukaemia; LP = lip; LU = lung; LV = liver; ME = skin, melanoma; MY = multiple myeloma; NHL = non-Hodgkin lymphoma; NO = nose; OE = oesophagus; OTH = other cancers; OV = ovary; PA = pancreas; RE = rectum; SK = skin, non-melanoma; SO = soft tissue; ST = stomach; TH = thyroid; TOT = all cancers combined; UR = urinary bladder.**

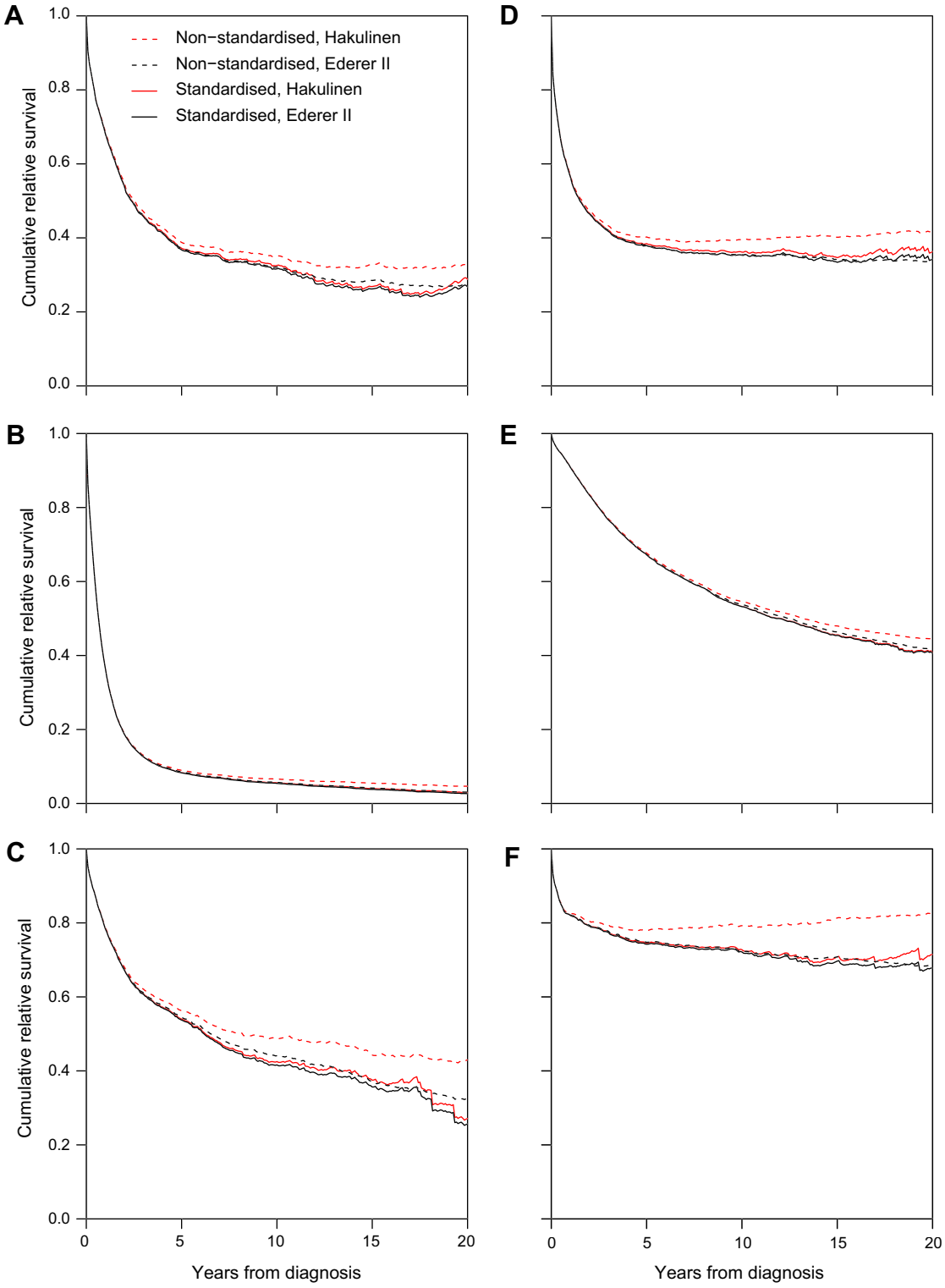


**Fig. 3 – Age-standardised and non-standardised cumulative relative survival ratios (%) of patients diagnosed in Finland in 1970–1979 and followed-up until the end of 2009 estimated using three different methods for cancers of the rectum (A), lung (B) and brain and central nervous system (C) in males and those of the colon (D), breast (E) and thyroid (F) in females. Results obtained by the Ederer I method (Ederer et al.[4]) equal to those given by the Hakulinen method (Hakulinen[6]). The Ederer II method was published in Ref. [3].**

numbers by age group. Cancers of the lip and nose in females are good examples on this in Fig. 2. In the most extreme situation, not all the age-specific relative survival estimates are available. However, even in this situation, the Ederer II method gives an estimate of cumulative relative survival.

The mathematics in Appendix A shows that when the excess hazards and thereby the conditional (annual) relative survival ratios become by follow-up time independent of age, the age standardisation is no longer important if the Ederer II method is utilised. This is the case with many cancers, with a particular special case of zero excess hazards or annual relative survival ratios of one across the ages, indicating patient cure.

The age-standardised cumulative relative survival ratios shown in Fig. 3 mostly agree with those derived with the Ederer II method. The largest differences are for cancers of the brain and central nervous system in males and thyroid cancer in females. The age-specific excess hazards approximately agree after five years of follow-up for the other cancer sites with the exception of lung cancer in males, for which as for cancer of the brain and thyroid, the excess hazards are higher for the older age groups (three of the sites shown as examples in Figs. 4–6).

The Ederer II method gives, by follow-up time, increasingly less weight for the excess hazard in high ages and more weight for those in the low ages than the internal age standardisation (Figs. 4–6). The differences between the two methods increase in that respect particularly when the cumulative relative survival for patients in high ages is high, as for female colon cancer in Fig. 4. If the cumulative relative survival is low for high ages as for lung cancer in Fig. 5, the differences are much smaller. The age-specific components of the overall excess hazard are for both colon and lung cancer roughly comparable between the two methods (Figs. 4 and 5). As a result, the overall excess hazards are fairly well comparable between the methods.

Cancer of brain and central nervous system for males shows the empirical extreme opposite result (Fig. 6). The excess hazards increase by age and, as a result of the substantial differences between the methods in the age-specific weights by follow-up time, the resulting age-specific components of the excess hazard are also different. However, the results on the overall excess hazards are still fairly comparable. As discussed in the Appendix A, usually the Ederer II method tends to slightly over-estimate the internally age-standardised

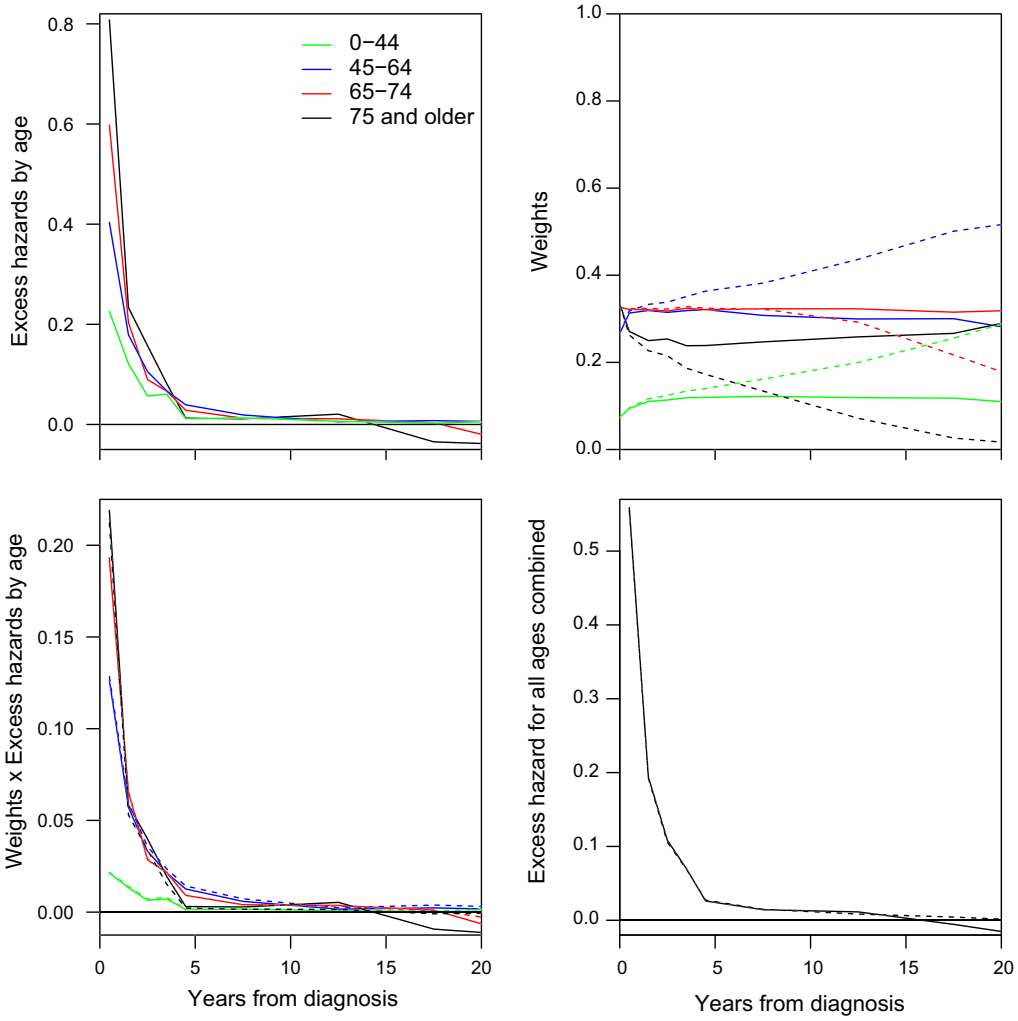


**Fig. 4 – Excess hazards, weights used for them and components of the overall excess hazards by age, and the overall excess hazards for the internal age standardisation and the Ederer II method (solid and broken lines, respectively), for female patients with cancer of the colon diagnosed in Finland in 1970–1979 and followed-up until the end of 2009. The excess hazards were calculated annually up to 5 years and thereafter for 5-year periods.**

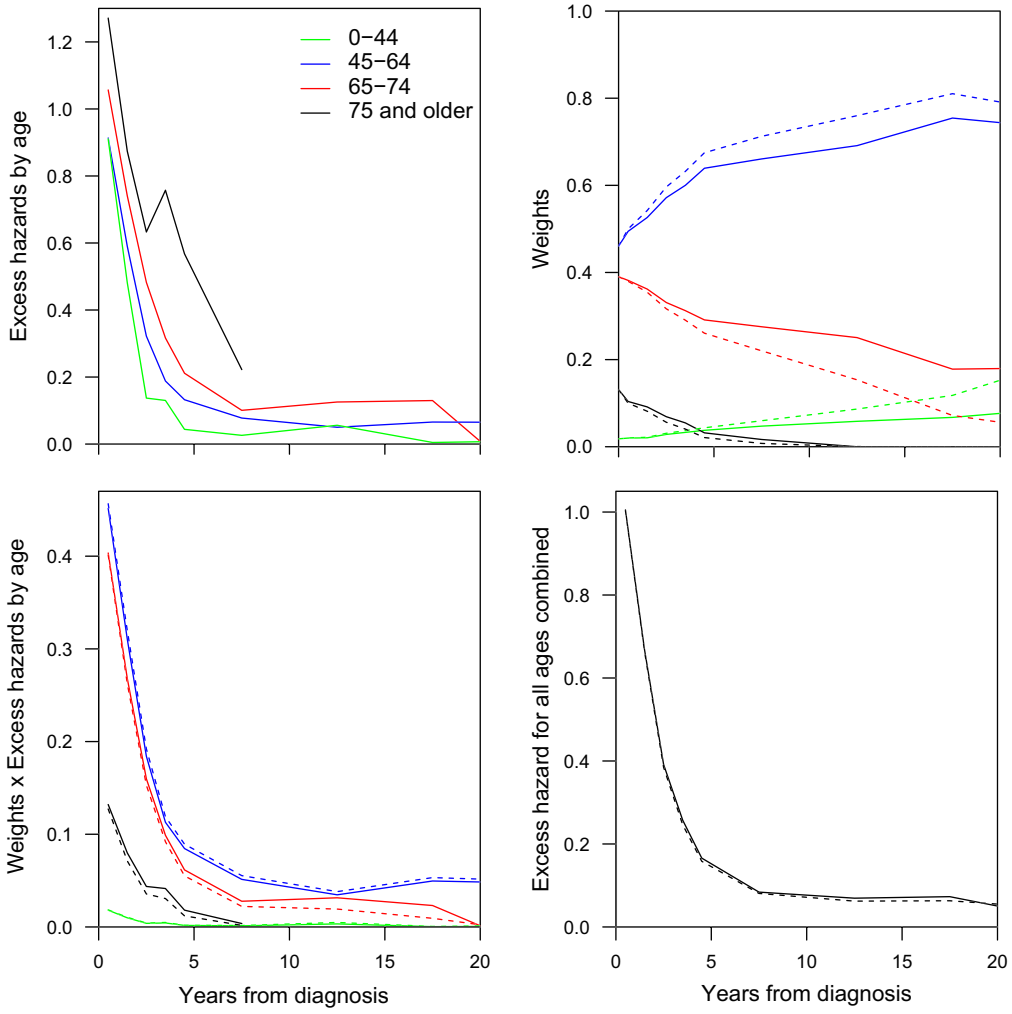


**Fig. 5 – Excess hazards, weights used for them and components of the overall excess hazards by age, and the overall excess hazards for the internal age standardisation and the Ederer II method (solid and broken lines, respectively), for male patients with cancer of the lung diagnosed in Finland in 1970–1979 and followed-up until the end of 2009. The excess hazards were calculated annually up to 5 years and thereafter for 5-year periods.**

relative survival, but by far to a lesser extent than the Hakulinen or Ederer I method (Figs. 1–3).

## 4. Discussion

The relative survival ratio was originally intended to give an estimate of the net survival when cancer of the patients was the only cause of death.[4] For that purpose, the internal age standardisation gives the gold standard. On the other hand, if the purpose is to compare the patients' survival with a theoretical maximum, survival to be obtained in the comparable general population group,[5] this gold standard does not apply.[7] For that purpose, the Ederer I method gives the gold standard if there is no censoring and the Hakulinen method when censoring is to be accounted for. In practical applications, estimation of the net probability of surviving for the patients is a more central goal than comparison of patients' survival against a theoretical maximum. Consequently, the Ederer II method is more often to be recommended than the two other methods.

The theory and empirical results favour the routine adoption of the Ederer II method[3] unless age standardisation is needed for external comparisons between groups. Use of the Ederer II method is particularly important for long-term follow-up where the estimation is more based on the younger age groups. This is also reflected in the precision of the cumulative estimates. In female colon-cancer patients, at 15 years, for example, the standard errors for the age-standardised ratio and for the ratio derived by the Ederer II approach are 1.9% units and 1.3% units, respectively. For lip cancer in females, the corresponding standard errors at the 15-year follow-up are 12.0 and 7.9% units for the two approaches.

There are drawbacks in the application of the Ederer II method related to informative censoring, and in principle it may appear controversial that the expected survival probability will depend on the observed survival in the previous intervals.[7] In practice, it is likely, however, that the effect of these problems will usually be small. In order to see what could happen in the circumstances of extreme differences in potential censoring between the age groups, let us return to the

original examples used when the Hakulinen method was introduced.[6] Let us use the two most extreme patterns of potential censoring:

$W_0$: 5% of the patients are admitted each year over a period of 20 years, and
$W_2$: 3.1% of the patients are admitted during the first year, and each year this figure is increased by 0.2% units (3.3%, 3.5%, . . ., 6.9%).

The common closing date of follow-up is at the end of the 20th year. For simplicity, it is assumed that the general population life tables are for the whole 20-year period as in 1971–1975 in Finland[8] and that the patients consist of two age categories, those aged 35 and 55 years at diagnosis, with age-specific relative survival equal to that calculated for ages 0–44 and 45–64 years, respectively, in Finland in 1953–1974.

Hakulinen[6] lacked a gold standard that can now be added (Table 1, the column labelled "Internal age standardisation"). Now, compared to the gold standard, the results of the Ederer II method are closer than those derived by the Ederer I or Haku-linen methods, particularly when the age-specific cumulative relative survival ratios differ and when the follow-up becomes longer (Table 1). The results of all the non-standardised methods depend on censoring, and particularly the Ederer I and Hakulinen methods tend to over-estimate the gold standard. In practice, the differences in the patterns of the potential censoring will be smaller than in this theoretical extreme example.

To make an internal standardisation or to use the Ederer II method makes a tradeoff between including random error or bias (systematic error) in the estimation. The Finnish Cancer Registry's materials are large and provide good empirical estimates on the bias resulting from the use of the Ederer II method. As a rule, this bias is rather small and leads to a small over-estimation. But even the Finnish Cancer Registry has small materials when the estimates based on internal standardisation become imprecise or impossible. In this situation, more stable estimates are derived by using the Ederer II method, at the cost of a possible small bias.

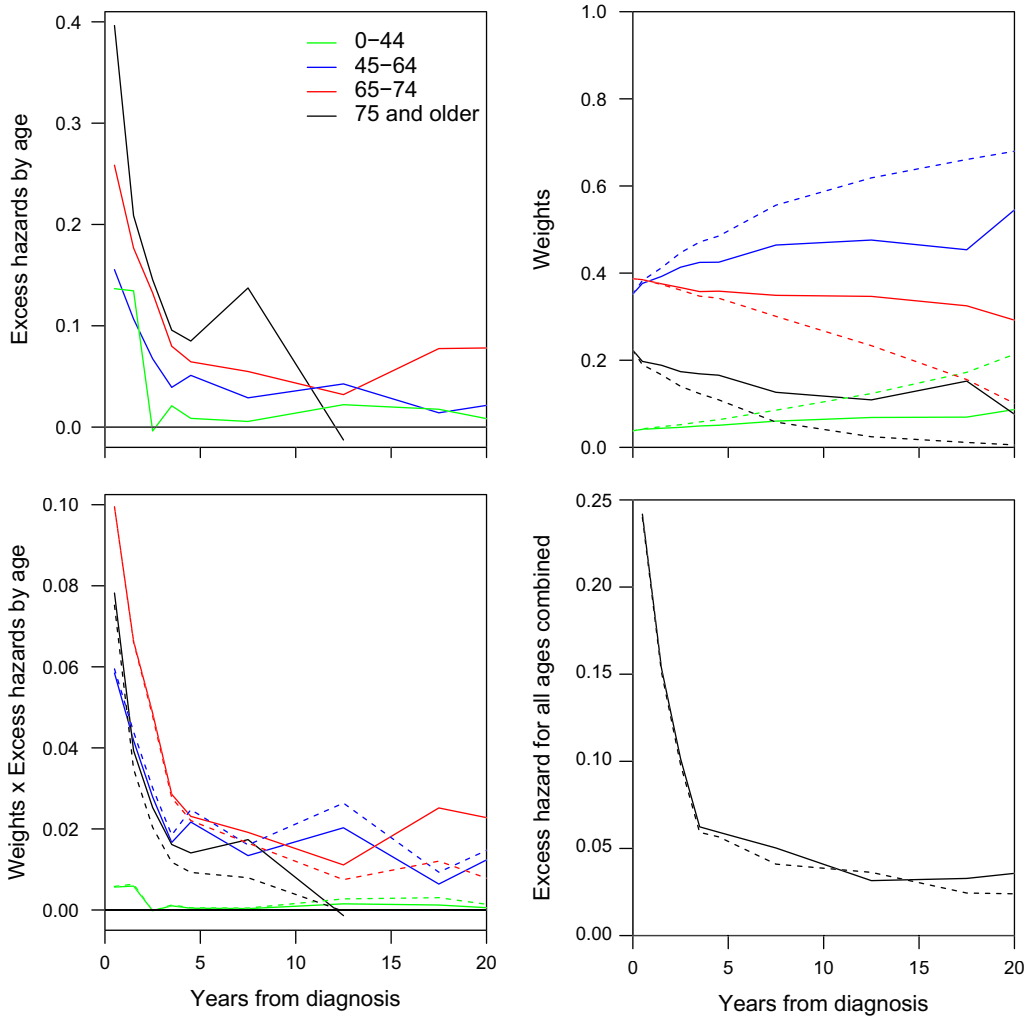


**Fig. 6 – Excess hazards, weights used for them and components of the overall excess hazards by age, and the overall excess hazards for the internal age standardisation and the Ederer II method (solid and broken lines, respectively), for male patients with cancer of the brain and central nervous system diagnosed in Finland in 1970–1979 and followed-up until the end of 2009. The excess hazards were calculated annually up to 5 years and thereafter for 5-year periods.**

**Table 1 – Long-term relative survival ratio (%) for cancer at three sites as assumed for patients aged 35 and 55 years, and as derived for the pooled group with four alternative methods, when the age-specific potential censoring patterns differ. Hypothetical example (see text).**

| Primary site and years of follow-up | Age | | Internal age standardisation | Method and potential censoring patterns for ages 35/55 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Ederer II | | | Ederer I | | | Hakulinen | |
| | 35 | 55 | | Homog.[a] | $W_2/W_0$ | $W_0/W_2$ | Homog.[a,b] | $W_2/W_0$ | $W_0/W_2$ | $W_2/W_0$ | $W_0/W_2$ |
| *Melanoma* | | | | | | | | | | | |
| 5 | 47.2 | 42.3 | 44.8 | 44.8 | 44.8 | 44.8 | 44.9 | 44.8 | 44.9 | 44.8 | 44.9 |
| 10 | 40.5 | 33.2 | 36.9 | 37.0 | 36.9 | 37.1 | 37.2 | 36.9 | 37.6 | 37.1 | 37.4 |
| 15 | 36.1 | 35.0 | 35.6 | 35.3 | 35.5 | 35.0 | 35.6 | 35.2 | 36.1 | 35.9 | 35.4 |
| 20 | 37.3 | 35.9 | 36.6 | 36.5 | 36.6 | 36.3 | 36.8 | 35.3 | 38.3 | 37.0 | 36.7 |
| *Thyroid* | | | | | | | | | | | |
| 5 | 78.5 | 50.2 | 64.4 | 64.9 | 64.4 | 64.7 | 64.9 | 64.7 | 65.1 | 64.8 | 65.1 |
| 10 | 77.7 | 44.1 | 60.9 | 61.3 | 61.0 | 61.7 | 62.7 | 62.0 | 63.3 | 62.4 | 63.0 |
| 15 | 71.3 | 25.1 | 48.2 | 50.0 | 48.5 | 51.4 | 52.8 | 50.3 | 55.0 | 51.3 | 54.0 |
| 20 | 59.8 | 19.6 | 39.7 | 41.5 | 40.0 | 42.8 | 46.3 | 43.1 | 49.1 | 45.2 | 47.0 |
| *Lung* | | | | | | | | | | | |
| 5 | 14.5 | 8.6 | 11.6 | 11.6 | 11.6 | 11.7 | 11.7 | 11.6 | 11.8 | 11.6 | 11.7 |
| 10 | 12.4 | 5.9 | 9.2 | 9.3 | 9.1 | 9.4 | 9.5 | 9.3 | 9.6 | 9.3 | 9.6 |
| 15 | 11.6 | 4.9 | 8.3 | 8.4 | 8.3 | 8.6 | 8.9 | 8.6 | 9.2 | 8.8 | 9.0 |
| 20 | 12.7 | 5.8 | 9.3 | 9.4 | 9.2 | 9.5 | 10.4 | 9.9 | 10.9 | 10.4 | 10.4 |

[a] Homogeneous censoring patterns.
[b] Also for the Hakulinen method.

Although, some mathematics has been given in Appendix A as a background, the results in this study on the small bias are empirical findings. It is easy to create non-realistic theoretical examples where the bias becomes large. Nevertheless, simulation studies can shed more systematically light on the question of the tradeoff between the random error and bias. These are, however, outside the scope of the present study.

Thus, in conclusion, there seems to be the time of making a reform in the reporting of cumulative relative survival and to adopt the method by Ederer and Heise.[3] However, age standardisation itself tends to suppress differences that may vary by age. Differences between groups of interest are often age dependent, and thus statistical modelling[9–11] will provide the best insight into the figures in national and even in international relative survival analyses.

## Conflict of interest statement

None declared.

## Sources of support

This work was supported by grants from the Cancer Society of Finland and the Academy of Finland (No. 122150).

## Appendix A. Some mathematics on the relative survival methods

### A.1. *Internal age standardisation (gold standard)*

The gold standard gives the hazard of dying and the probability of surviving when the hazard of the patients from dying of their cancer is the only mortality hazard. Let us assume that this hazard at time $t$, $v_a(t)$, varies by age category $a$. Then, the probability of surviving up to $t$ in age group $a$ is

$$r_a(t) = \exp\left[-\int_0^t v_a(u)\,du\right].$$

When the mortality hazard $v_a(t)$ due to the cancer of the patients is defined as the difference between the total mortality hazard $\mu_a(t)$ of the patients and the expected hazard $\mu_a^*(t)$ in the comparable general population group, the difference is called the excess hazard due to cancer. Converting this to the probability scale gives the relative survival ratio.

For all ages combined, the corresponding probability is

$$r_s(t) = \sum_a w_a r_a(t),$$

in which $w_a$ is the proportion of patients belonging to $a$ at $t = 0$. This can also be regarded as an age-standardised relative survival ratio with internal weights $w_a$. The excess hazard for all ages combined, is then

$$v_s(t) = \sum_a \frac{w_a r_a(t) v_a(t)}{\sum_a w_a r_a(t)}$$

and

$$r_s(t) = \exp\left[-\int_0^t v_s(u)\,du\right].$$

If all the excess hazards by age group are equal, $v_a(t) = v(t)$ for all $a$, also $v_s(t) = v(t)$.

### A.2. *Method by Ederer et al. (1961) (Ederer I)*

Let us assume that there is no informative censoring by age and $p_a(t)$ = observed cumulative survival probability for age

group $a$ up to $t$. By writing $W_a(t) = w_a p_a(t)$, the observed mortality hazard at $t$

$$\mu(t) = \sum_a \frac{W_a(t)\mu_a(t)}{\sum_a W_a(t)},$$

is a weighted average of the age-specific hazards at $t$, $\mu_a(t)$.

Let $p^*_a(t)$ = cumulative expected survival probability in the general population for age group $a$ up to $t$. By writing $W^*_a(t) = w_a p^*_a(t)$, the expected mortality hazard at $t$

$$\mu^*_I(t) = \sum_a \frac{W^*_a(t)\mu^*_a(t)}{\sum_a W^*_a(t)},$$

is a weighted average of the age-specific expected hazards at $t$, $\mu^*_a(t)$. The weights differ from those of the observed age-specific hazards.

The excess hazard at $t$ is, by writing $\mu_a(t) = \mu^*_a(t) + v_a(t)$,

$$\begin{aligned} v_I(t) &= \mu(t) - \mu^*_I(t) \\ &= \sum_a \frac{W_a(t)\mu^*_a(t)}{\sum_a W_a(t)} - \sum_a \frac{W^*_a(t)\mu^*_a(t)}{\sum_a W^*_a(t)} + \sum_a \frac{W_a(t)v_a(t)}{\sum_a W_a(t)}. \end{aligned}$$

If $r_a(t) = r(t)$ and $v_a(t) = v(t)$ for all $a$, $p_a(t) = p^*_a(t)r_a(t) = p^*_a(t)r(t)$ and $W_a(t)/\sum_a W_a(t) = W^*_a(t)/\sum_a W^*_a(t)$ and thereby $v_I(t) = v(t)$.

Often, however, the $r_a(t)$ are lower for the higher age groups $a$. The $\mu^*_a(t)$ are always higher for the higher age groups. Thus, the weights $W_a(t)/\sum_a W_a(t)$ tend to give less weight to the $\mu^*_a(t)$ of high ages than the $W^*_a(t)/\sum_a W^*_a(t)$. If these differences exist, it follows that even when $v_a(t) = v(t)$ for all $a$ at $t$, $v_I(t) < v(t)$. Thus, differences in cumulative relative survival that have emerged before $t$ have an impact on the excess hazard at $t$, and usually the result is an under-estimation of the excess hazard. The cumulative relative survival

$$r_I(t) = \exp\left[-\int_0^t v_I(u)\,du\right]$$

then over-estimates the $r_s(t)$.

### A.3. *Method by Ederer and Heise (1959) (Ederer II)*

Let us first look at hazards in the absence of informative censoring. The expected hazard at $t$ is then

$$\mu^*_{II}(t) = \sum_a \frac{W_a(t)\mu^*_a(t)}{\sum_a W_a(t)},$$

and the excess hazard

$$v_{II}(t) = \mu(t) - \mu^*_{II}(t) = \sum_a \frac{W_a(t)[\mu_a(t) - \mu^*_a(t)]}{\sum_a W_a(t)} = \sum_a \frac{W_a(t)v_a(t)}{\sum_a W_a(t)}.$$

If $v_a(t) = v(t)$ for all $a$, also $v_{II}(t) = v(t)$. Note that, unlike for the Ederer I method, no assumption on equality of the age-specific cumulative relative survival ratios is needed.

For informative censoring, let $c_a(t)$ = proportion of patients in age group $a$ with a potential follow-up until $t$. The equation above holds for $v_{II}(t)$ also in this case by writing $W_a(t) = w_a p_a(t)c_a(t)$. It is easy to see that if $v_a(t) = v(t)$ for all $a$, $v_{II}(t) = v(t)$ holds even under informative censoring. Thus, the cumulative relative survival ratio $r_{II}(t) = \exp[-\int_0^t v_{II}(u)\,du]$ is multiplicatively incremented in this situation exactly as the gold standard $r_s(t)$, with increments of the form $\exp[-\int_{t_1}^{t_2} v(u)\,du]$.

In general, however, when the age-specific excess hazards are not equal, the weights in $v_s(t)$, $w_a r_a(t)/\sum_a w_a r_a(t)$ give more weight to the excess hazards of the high age groups $v_a(t)$ than the weights $W_a(t)/\sum_a W_a(t)$ due to inclusion of $p^*_a(t)$ in $W_a(t) = w_a p^*_a(t) r_a(t)$. Thus, in general, even the Ederer II method may over-estimate the gold standard $r_s(t)$ when the age-specific excess hazards are unequal. Informative censoring may, however, distort this general tendency.

### A.4. *Hakulinen method (1982)*

This is in principle the same as the Ederer I method but accounting for informative censoring. The mathematics is derived simply by inserting $w_a(t) = w_a p_a(t) c_a(t)$ and $W^*_a(t) = w_a p^*_a(t) c_a(t)$ into the Ederer I method expressions. It is easy to see that if $r_a(t) = r(t)$ and $v_a(t) = v(t)$ for all $a$, $v_I(t) = v(t)$ also under informative censoring. Informative censoring may distort the tendency of over-estimation of cumulative relative survival when the $r_a(t)$ are not equal.

## Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejca.2011.03.011.

## REFERENCES


1. Ries LAG, Melbert D, Krapcho M, et al., editors. *SEER cancer statistics review, 1975–2004*. Bethesda, MD: National Cancer Institute; 2007.
2. Coleman MP, Quaresma M, Berrino F, et al. Cancer survival in five continents: a worldwide population-based study (CONCORD). *Lancet Oncol* 2008;**9**:730–56.
3. Ederer F, Heise H. *Instructions to IBM 650 programmers in processing survival computations. Methodological note no. 10*. Bethesda, MD: End Results Evaluation Section, National Cancer Institute; 1959.
4. Ederer F, Axtell LM, Cutler SJ. *The relative survival rate: A statistical methodology. Monograph no. 6*. Bethesda, MD: National Cancer Institute; 1961.
5. Hakulinen T. On long-term relative survival rates. *J Chron Dis* 1977;**30**:431–43.
6. Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics* 1982;**38**:933–42.
7. Pokhrel A, Hakulinen T. How to interpret the relative survival ratios of cancer patients. *Eur J Cancer* 2008;**44**:2661–7.
8. Central Statistical Office of Finland. Mortality. Life Tables 1971–75. Official Statistics of Finland VI A: 142. Helsinki: Central Statistical Office of Finland.
9. Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Stat Med* 2004;**23**:51–64.
10. Nelson CP, Lambert PC, Squire IB, Jones DR. Flexible parametric models for relative survival, with application in coronary heart disease. *Stat Med* 2007;**26**:5486–98.
11. Remontet L, Bossard N, Belot A, Estève J. FRANCIM. An overall strategy based on regression models to estimate relative survival and model the effects of prognostic factors in cancer survival studies. *Stat Med* 2007;**26**:2214–28.