PERGAMON

European Journal of Cancer 38 (2002) 690–695

European Journal of Cancer

www.ejconline.com

A computer program for period analysis of cancer patient survival

H. Brenner[a,*], O. Gefeller[b], T. 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-91054 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 31 July 2001; received in revised form 14 November 2001; accepted 19 December 2001

Abstract

Monitoring of long-term survival rates, which is now routinely performed by many cancer registries throughout the world, should be as up-to-date as possible. A few years ago, a new method of survival analysis, denoted period analysis, has been proposed which provides more up-to-date estimates of long-term survival rates than traditional survival analysis by exclusively reflecting the survival experience of patients within a recent calendar period. However, application of this method has so far been hindered by the lack of pertinent computer programs. In this paper, we present a simple and easy-to-use computer program (SAS macro) that enables one to carry out period analysis (as well as conventional analysis) of both absolute and relative survival rates with the type of data commonly available in population-based cancer registries. We illustrate application of the program with examples from the nationwide Finnish Cancer Registry. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Cancer registries; Statistical methods; Survival

1. Introduction

Long-term survival rates are essential outcome measures of patients with cancer. Monitoring of long-term survival rates is now routinely performed by many cancer registries throughout the world. The most commonly reported measures of long-term survival reported by population-based cancer registries are 5- and 10-year survival rates (e.g. Refs. [1–7]). Occasionally, survival rates are also reported for longer follow-up times, such as 15 or 20 years [8].

Survival rates are typically monitored and reported for patients diagnosed within defined calendar years by traditional methods of survival analysis [9–12]. Often, only patients who have been followed over the full follow-up period of interest are included in the analysis. For example, the most recent estimate of 10-year survival of patients with acute lymphocytic leukaemia provided in a recent report from England and Wales pertained to patients diagnosed between 1981 and 1985 and followed until the end of 1995 [13].

As survival estimates derived in this way of a cohort approach pertain to survival experience of patients diagnosed many years ago, they do not reflect potential recent progress in prognosis, and they may strongly lag behind the survival expectations of newly diagnosed patients in cases of recent improvement in prognosis. Other analyses also include more recently diagnosed patients who have not completed the full follow-up period of interest at the closing date of follow-up, but who are censored at that point of time. For example, a recent survival analysis from the Finnish Cancer Registry included patients diagnosed in 1985–1994 and followed until the end of 1995 to derive the most up-to-date 10-year survival rates [14]. However, even with this approach, the most up-to-date long-term survival estimate reflects survival experience over a time period that dates back quite a long term in the past, and it may still be quite outdated in case of recent improvement of survival.

A few years ago, a new method of survival analysis, denoted period analysis, has been proposed in order to obtain more up-to-date estimates of long-term survival

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

0959-8049/02/$ - see front matter
PII: S0959-8049(02)00003-5

rates [15,16]. Period analysis exclusively reflects the survival experience of patients within a most recent calendar period, e.g. the most recent calendar year, for which mortality follow-up is available. This is achieved by left truncation of observations at the beginning of this period, in addition to right censoring at its end as previously described in Refs. [15,16].

Recently, period analysis has been used by various cancer registries to derive more up-to-date estimates of long-term survival rates [17–21]. However, more widespread application has been hindered by the lack of pertinent computer programs which is often the obstacle retarding the wider utilisation of a new statistical method. In this paper, we present a simple and easy-to-use computer program that enables one to carry out period analysis of survival with the type of data commonly available in population-based cancer registries, and we illustrate its application by examples from the Finnish Cancer Registry. We thereby hope to facilitate use of this new approach by cancer registries interested in deriving more up-to-date estimates of long-term survival rates.

2. The Computer Program

2.1. Statistical method

The program uses a life table approach in which conditional survival probabilities for 1-year intervals following diagnosis are combined to derive cumulative survival rates. Both absolute and relative survival rates can be estimated. The relative survival rates, which are commonly reported by cancer registries along with absolute (observed) survival rates, reflect the net mortality of patients in the hypothetical situation in which cancer is the only cause of death. The relative survival rates are derived as the ratios of the observed survival rates to the expected rates for groups of people in the general population similar to the patient group with respect to sex, age and calendar period of observation [11,12].

In our program, the relative survival rates are estimated according to the method commonly known as the Ederer II method [22] adapted (by left truncation of observations at the beginning of the period of interest) to the period analysis approach. Point estimates and standard errors, calculated according to Greenwood's method [23], are provided.

2.2. Software requirements

The program is written as a SAS macro called '*period*' and can be run with the SAS statistical software package version 8 or older. The macro and its documentation can be downloaded free of charge from the statistical archive network maintained by the Department of Informatics, Biometry and Epidemiology at the University of Erlangen-Nuremberg (http://www.imbe.med.uni-erlangen.de/issan/issan.html). Extensive validation efforts have been made to guarantee the correctness of the macro. The macro is presented and explained in detail below, so that investigators may easily adapt the program to their specific needs.

2.3. Data requirements

As with traditional survival analysis, the following minimal information is needed for each patient included in the analysis (notation and definition of variables in parentheses; see also Table 1): Month of diagnosis (*dm*, $1 \leqslant dm \leqslant 12$) and year of diagnosis (*dy*, four digits, e.g. 1990), month of end of follow-up (*fm*, $1 \leqslant fm \leqslant 12$) and year of end of follow-up (*fy*, four digits, e.g. 1995), and vital status at the end of follow-up (*vitstat*; 1 = alive, 2 = dead). Patients with missing information on any of these variables should be excluded prior to analysis. Although the program may easily be extended to allow for the use of more precise dates (e.g. day of diagnosis and death), this is neither necessary nor commonly done for most applications of monitoring of cancer patient survival.

If relative survival rates are to be obtained in addition to absolute survival rates (or if age- and sex-specific analyses are performed), which is common practice in the cancer registry world, one also needs to have sex (*sex*; 1 = males, 2 = females) and age at diagnosis (*diagage*; in years) of each patient, as well as conditional 1-year survival rates by age and sex of the general population for the calendar period included in the analysis. This information is commonly provided as an integral component of period life tables by national or state statistical offices. In the absence of such life tables, they may also be easily be constructed from age- and sex-specific mortality rates as described in detail

Table 1
List of variables needed for each patient and required specifications

Variables needed for all patients	
diagage	age at diagnosis (in years)
sex	gender (1 = males, 2 = females)
dy	year of diagnosis (e.g. 1990)
dm	month of diagnosis (e.g. 6 for June)
fy	year of end of follow-up (e.g. 1995)
fm	month of end of follow-up (e.g. 12 for December)
vitstat	vital status at the end of follow-up (1 = alive, 2 = dead)
Specifications	
k	length of follow-up (e.g. $k = 10$ for 10-year survival rates)
perbeg	first calendar year of period of interest (e.g. 1995)
perend	last calendar year of period of interest (e.g. 1997)
P[2,100]	2×100 dimensional array Elements P [*sex*,*toage*] indicate conditional survival rates from age *toage-1* to age *toage* of male (*sex* = 1) or female (*sex* = 2) subjects in the general population during the period of interest

elsewhere in [24]. The conditional one year survival rates by sex and age must be read in a 2×100 dimensional array (*P[2,100]*) with elements *p*[*sex,toage*], where *toage* = 1,. . .,100 indicates the upper end of 1-year age intervals. For example, *p[1,50]* equals the conditional 1-year survival of a male person from his 49th to his 50th birthday.

Other specifications that need to be made include the length of follow-up in years (denoted k, e.g. $k=10$ for 10-year survival rates), and the first and last calendar year of the period of interest (denoted *perbeg* and *perend*, respectively). These settings can be specified in the macro call.

2.4. Macro composition

The macro *period*, which is shown in the Appendix, includes four steps. Each of these steps is outlined by pertinent comments (lines starting with '**') in the macro.

Step 1, which is the core step of the macro, determines the contribution of each patient to person-times at risk (*perl[i]*), observed deaths (*perd[i]*), as well as expected deaths in the absence of cancer (*perde[i]*) for each follow-up year i ($1<=i<=k$). This step starts with specifications of pertinent arrays of length k, whose cells are initially set to 0. All further calculations are done iteratively for each calendar year (*cy*) included in the period of interest, and they are restricted to those patients whose year of diagnosis equals or precedes the *cy* by up to k years, and who are still at risk in the *cy* (only those patients may contribute deaths and/or person-time at risk in *cy*). For these patients, their age in *cy* is calculated (*age*, in years), and the expected 1-year probability of survival of a person of the same age and sex in the general population, denoted *gp*, is derived from the population life table. Because life tables are typically limited to ages below 100 years, and because the number of subjects above that age is typically very small, the simplifying assumption is made that the 1-year survival probability of subjects above the age of 100 years in the general population equals the 1-year survival probability of 99-year-old people (this assumption has very little impact on the overall estimates of relative survival).

The contribution of each patient to the person-time at risk and to the observed and expected number of deaths are then determined for each follow-up year i following diagnosis ($1\leqslant i\leqslant k$). Follow-up year specific person-times at risk and observed and expected numbers of death are summed up over calendar years *cy* included in the period of interest.

In step 2 of the macro, the follow-up year specific person-times at risk and observed and expected numbers of deaths are summed up over all patients. These sums (*sperl[i]*, *sperd[i]* and *sperde[i]*, $1\leqslant i\leqslant k$) are written into a new dataset called *sumvar*.

In step 3, cumulative absolute and relative survival rates (*abssur[i]* and *relsur[i]*, respectively) with their standard errors (*abserr[i]* and *relerr[i]*, respectively) are derived for each follow-up year i ($1\leqslant i\leqslant k$) using standard methodology [12].

In step 4, the requested output is specified. The default ouput consists of the cumulative absolute and relative survival rates and their standard errors for each year (up to the kth year) following diagnosis, but this may be easily customised as desired by simple modifications of the variables included in the 'print' procedure (second line from bottom).

2.5. Use of the program for conventional survival analysis

The program may also be used for conventional survival analysis by simply including all calendar years from the earliest year of diagnosis of the patients to the last calendar year of the follow-up (to be defined by appropriate specifications for *perbeg* and *perend*) in the period of interest. If only patients diagnosed within defined calendar years are to be included, the sample needs to be restricted accordingly prior to invoking the macro.

Calculations of long-term survival rates by traditional methods typically include survival experience of patients over a wide range of calendar years, and it may be prudent to refer to different life tables for the general population rather than to a single life table for different calendar years in the derivation of relative survival rates. This may easily be accomplished by additional invoking of a macro that reads the life table information for the specific calendar year of interest into the array *P[2,100]* at the beginning of the calendar year specific calculations in step 1 of the macro *period* (such an addition may also be considered for period analyses encompassing calendar periods of more than one calendar year).

3. Empirical examples

We illustrate use of the macro with an analysis intended to derive up-to-date estimates of 10-year relative survival rates of patients with various forms of cancer in Finland. Our estimates are based on data of the nationwide Finnish Cancer Registry, which is among the highest quality cancer registries in the world. Virtually complete cancer registration and mortality follow-up has been achieved since the early 1950s [25,26]. At the time of this analysis, the data base of cancer registration and mortality follow-up had been completed until the end of 1997.

We calculated 10-year absolute and relative survival estimates for various common forms of cancer (colon,

lung, breast, prostate) and their standard errors for the periods 1990–1992, 1993–1995 and 1996–1997 to investigate recent trends in long-term prognosis. We *a priori* restricted the sample to patients diagnosed in 1980 or later years, because patients diagnosed in earlier years have completed all of their 10-year follow-up prior to 1990. We used a period life table for 1991–1995 in all analyses. The macro calls for the three analyses are simply given by %period (10, 1990, 1992), %period (10, 1993, 1995) and %period (10, 1996, 1997), respectively.

The results are presented in Table 2. As expected, prognosis varied according to the type of cancer. The highest 10-year survival rates were seen for breast cancer, and the lowest 10-year survival rates were seen for lung cancer. The difference between relative and absolute survival rates was largest for prostate cancer, as the mean age and hence mortality from other causes is much higher for patients with this cancer than for other patients. For all types of cancer, except for lung cancer, our analysis shows substantial recent improvement in long-term survival rates. In the 1996–1997 period, the 10-year relative survival rate exceeded 55% for both colon and prostate cancers, and 75% for breast cancers. These improvements had not been disclosed in previous reports which were based on traditional methods of survival analysis [14].

4. Discussion

In this paper, we provide an easy-to-use computer program (SAS macro) that should facilitate use of the period approach for estimating long-term absolute and relative survival rates, which are commonly reported by cancer registries for monitoring progress against cancer. The period approach which has been introduced some years ago leads to more up-to-date estimates of long term survival rates than traditional methods of survival analysis [15,16]. Nevertheless, even period estimates may still lag behind the survival expectations of newly diagnosed patients in the case of ongoing improvement in survival. In theory, period analysis might also provide overoptimistic estimates of long-term survival rates in situations in which recent advancements in cancer care or early diagnosis merely postpone death rather than increase the chance of survival. However, an extensive empirical evaluation of the method in the database of the Finnish Cancer Registry, which encompasses registrations over almost half a century, revealed that this theoretical possibility is of much less concern in practice than potential underestimation of survival rates, even though the latter is typically much less severe for period analysis than for traditional survival analysis [27,28]. Like other forms of survival analysis, period analysis does not by itself disclose the reasons for increased survival. They may include earlier diagnosis (in which case survival time may only be increased by the lead time in some instances) or more effective therapy. In the unusual case of deteriorating survival rates over time, such an alarming development would likewise be more timely disclosed by period analysis than by traditional methods of survival analysis.

In contrast to survival analysis, the principle of period analysis is well established for other health statistics. For example, up-to-date estimates of life expectancy are commonly derived from period life tables, which reflect mortality rates of people at various ages (belonging to various birth cohorts) during a recent calendar period, such as a recent calendar year, rather than from cohort life tables, which would reflect mortality rates of birth cohorts born at least a life-span ago. Similarly, cumulative rates or risks up to a certain age, e.g. age 75 years, which are well established measures of cancer incidence or mortality [26], are derived for the most recent calendar periods for which age-specific data on cancer incidence are available rather than for cohorts of patients born 75 years ago (the latter would not even be possible with existing cancer registries).

The program presented in this paper uses a modified life table approach, in which cumulative long-term survival rates are derived from conditional survival probabilities within 1-year intervals following diagnosis. Like conventional life table analysis, this approach is based on the assumptions that censoring of patients is independent from their risk of death, and that the probability of death is constant within time intervals. Although the latter assumption does typically not fully hold in practice, the impact of its violation on the estimation of long-term survival rates is typically negligible for practical purposes, and life table analyses using 1-year time intervals are common practice in the cancer registry world [8,12]. The program may be adapted to allow for shorter time intervals, but the (typically very modest) gain in accuracy may, for practical purposes, not be worth the increased complexity this would imply.

Table 2
Period estimates of absolute and relative 10-year survival rates (in %; standard errors in parentheses) for common forms of cancer in Finland in 1990–1992, 1993–1995 and 1996–1997

Cancer site		1990–1992	1993–1995	1996–1997
Colon	Absolute	28.2 (0.9)	29.7 (0.9)	33.4 (1.1)
	Relative	46.8 (1.5)	49.8 (1.5)	57.0 (1.8)
Lung	Absolute	4.4 (0.2)	4.4 (0.3)	5.0 (0.4)
	Relative	6.5 (0.4)	6.6 (0.4)	7.5 (0.5)
Breast	Absolute	51.4 (0.6)	54.0 (0.6)	58.5 (0.7)
	Relative	66.5 (0.8)	69.9 (0.8)	75.2 (0.9)
Prostate	Absolute	18.1 (0.7)	19.1 (0.7)	23.0 (0.9)
	Relative	43.2 (1.6)	46.5 (1.6)	56.1 (2.1)

As it stands, the output of the program is limited to a list of absolute and relative cumulative survival rates and their standard errors. The program may be easily modified to provide additional output, such as the numbers of deaths and person-times at risk, as well as the conditional survival probabilities by follow-up interval. Other possible extensions might include further processing of the results, e.g. for generating graphical presentations of survival curves, but this is beyond the scope of this paper.

The program is provided as a SAS macro, and it should therefore be readily applicable by scientists who have access to this widely used software package. For those who use other statistical software packages, it should not be too difficult to translate the macro accordingly, given its brevity and simple structure. Apart from its use for period analysis of survival, the program may also facilitate carrying out conventional (cohortwise) analyses of relative survival rates (as described above), as procedures for estimating relative survival rates are typically not included in commonly used commercial software packages and therefore require special software.

In conclusion, we hope that our paper may be helpful to researchers interested in conducting period analyses of survival rates, and that it may promote the use of this approach for deriving more up-to-date estimates of long-term survival of patients with cancer.

Acknowledgement

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

Appendix. Macro 'period'

```
%macro period(k,perbeg,perend);

** Step 1: Determining the contribution of each patient
           to person-times at risk (perl[i]), observed
           deaths (perd[i]), and expected deaths in the
           absence of cancer (perde[i]) for each follow-
           up year i (1 < =i < =k);

array perl [&k]; array perd [&k]; array perde [&k];
do i=1 to &k; perl[i]=0; perd[i]=0; perde[i]=0; end;

  do cy=&perbeg to &perend;
      i=cy-dy;
      if 0 < =i < =&k and cy < =fy then do;
          age=diagage+(cy-dy);
          if age>99 then gp=p[sex,100];
              else if age=0 then gp=p[sex,1];
                  else gp=(p[sex,age]+p[sex,age+1])/2;
          ** note: gp=expected 1-year survival prob-
          ability of a person of the same sex and age in
          the general population;
          if i>0 then do;
              perl[i]=perl[i]+0.5;    perde[i]=perde[i]+
              0.5*(1-gp); end;
          if ((vitstat=1 and fm> =dm) or cy<fy) and
          i<&k then do;
              perl[i+1]=perl[i+1]+0.5;   perde[i+1]=
              perde[i+1]+0.5*(1-gp); end;
          if vitstat=2 and fm<dm and i>0 and fy=cy
          then perd[i]=1;
          if vitstat=2 and fm> =dm and fy=cy and
          i<&k then do;
              perl[i+1]=1; perde[i+1]=1-gp;
              perd[i+1]=1; end;
      end;
  end;

** Step 2: Summing up person times at risk, observed
           and expected deaths over all patients
           (dataset sumvar, variables sperl[i], sperd[i]
           and sperde[i], respectively);

proc means noprint sum; var perl1-perl&k perd1-
perd&k perde1-perde&k;
output out=sumvar sum=sperl1-sperl&k sperd1-
sperd&k sperde1-sperde&k; run;

** Step 3: Determining absolute cumulative survival
           rates (abssur[i]) and their standard errors
           (abserr[i]), expected cumulative survival
           rates in the absence of cancer (expsur[i]), and
           relative cumulative survival rates (relsur[i])
           and their standard errors (relerr[i]) for each
           follow-up year i (1 < =i < =k);

data sumvar; set sumvar;
array sperl[&k]; array sperd[&k]; array sperde[&k];
array abssur[&k]; array expsur[&k];
array vari[&k]; array abserr[&k]; array relsur[&k]; array
relerr[&k];

  do i=1 to &k;
      if sperd[i]>0 and sperd[i]> =sperl[i] then
      abssur[i]=0;
          else if i=1 then abssur[i]=1-sperd[i]/sperl[i];
              else if abssur[i-1]=0 then abssur[i]=0;
                  else abssur[i]=abssur[i-1]*(1-sperd[i]/
                  sperl[i]);
      if i=1 then expsur[i]=1-sperde[i]/sperl[i];
          else expsur[i]=expsur[i-1]*  (1-sperde[i]/
          sperl[i]);
      if abssur[i]=0 then vari[i]=.;
          else if i=1 then vari[i]=sperd[i]/(sperl[i]*
          (sperl[i]-sperd[i]));
```

```
            else     vari[i] = vari[i-1] + sperd[i]/(sperl[i]*
            (sperl[i]-sperd[i]));
        ** note: var[i] is the variance of absolute cumu-
        lative survival for follow-up year i;
        abserr[i] = abssur[i]*sqrt(vari[i]);
        if abssur[i] = 0 then relsur[i] = 0;
            else relsur[i] = abssur[i]/expsur[i];
        relerr[i] = abserr[i]/expsur[i];
    end;


** Step 4:  specification of output;

proc print; var abssur1-abssur&k abserr1-abserr&k
relsur1-relsur&k relerr1-relerr&k;
%mend period;
```

References

1. Berrino F, Sant M, Verdecchia A, Capocaccia R, Hakulinen T, Estève J, eds. *Survival of Cancer Patients in Europe: The EUROCARE Study*. *IARC Scientific Publication No. 132*. Lyon, IARC, 1995.
2. Berrino F, Capocaccia R, Estève J, *et al*., eds. *Survival of Cancer Patients in Europe: The EUROCARE-2 Study*. *IARC Scientific Publication No. 151*. Lyon, IARC, 1999.
3. Landis SH, Murray T, Bolden S, Wingo PA. Cancer statistics, 1999. *CA Cancer J Clin* 1999, **49**, 8–31.
4. Greenlee RT, Murray T, Bolden S, Wingo PA. Cancer statistics, 2000. *CA Cancer J Clin* 2000, **50**, 7–33.
5. Sankaranarayanan R, Black RJ, Parkin DM. *Cancer Survival in Developing Countries*. *IARC Scientific Publication No. 145*. Lyon, IARC, 1999.
6. Coebergh JW, Capocaccia R, Gatta G, Magnani C, Stiller CA. Childhood cancer survival in Europe, 1978–1992: the EUROCARE study. *Eur J Cancer* 2001, **37**, 671–672.
7. Terracini B, Coebergh JW, Gatta G, *et al*. Childhood cancer survival in Europe: an overview. *Eur J Cancer* 2001, **37**, 810–816.
8. Wingo PA, Gloeckler Ries LA, Parker SL, Heath CW. Long-term cancer patient survival in the United States. *Cancer Epidemiol Biomark Prev* 1998, **7**, 271–282.
9. Cutler SJ, Ederer F. Maximum utilization of the life table method in analyzing survival. *J Chron Dis* 1958, **8**, 699–712.
10. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958, **53**, 457–464.
11. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. *Natl Cancer Inst Monogr* 1961, **6**, 101–121.
12. Parkin DM, Hakulinen T. Analysis of survival. In Jensen OM, Parkin DM, Maclennan R, Muir CS, Skeet RG, eds. *Cancer Registration. Principles and Methods. IARC Scientific Publication No. 95*. Lyon, IARC, 1991, 159–176.
13. Schillinger JA, Grosclaude PC, Honjo S, Quinn MJ, Slogett A, Coleman MP. Survival after acute lymphocytic leukaemia: effects of socioeconomic status and geographic region. *Arch Dis Child* 1999, **80**, 311–317.
14. Dickman PW, Hakulinen T, Luostarinen T, *et al*. Survival of cancer patients in Finland. *Acta Oncol* 1999, **38**(Suppl. 12), 1–103.
15. Brenner H, Gefeller O. An alternative approach to monitoring cancer patient survival. *Cancer* 1996, **78**, 2004–2010.
16. Brenner H, Gefeller O. Deriving more up-to-date estimates of long term patient survival. *J Clin Epidemiol* 1997, **50**, 211–216.
17. Brenner H, Stegmaier C, Ziegler H. Recent improvement in survival of breast cancer patients in Saarland/Germany. *Br J Cancer* 1998, **78**, 694–697.
18. Brenner H, Stegmaier C, Ziegler H. Trends in survival of patients with ovarian cancer in Saarland, Germany, 1976–1995. *J Cancer Res Clin Oncol* 1999, **125**, 109–113.
19. Brenner H, Gefeller O, Stegmaier C, Ziegler H. More up-to-date monitoring of long-term survival rates by cancer registries: an empirical example. *Meth Inform Med* 2001, **40**, 248–252.
20. Brenner H, Kaatsch P, Burkhardt-Hammer T, Harms DÖ, Schrappe M, Michaelis J. Long term survival of children with leukemia achieved by the end of the 2nd millenium. *Cancer* 2001, **92**, 1977–1983.
21. Brenner H, Hakulinen T. Long-term cancer patient survival achieved by the end of the 20th century. Most up-to-date estimates from the nationwide Finnish cancer registry. *Brit J Cancer* 2001, **85**, 367–371.
22. Ederer F, Heise H. *Instructions to IBM 650 Programmers in Processing Survival Computations. Methodological Note No. 10, End Results Evaluation Section*. Bethesda, MD, National Cancer Institute, 1959.
23. Greenwood M. *A Report on The Natural Duration of Cancer*. London, Ministry of Health, His Majesty's Stationery Office, 1926.
24. Estève J, Benhamou E, Raymond L. *Statistical Methods in Cancer Research, Volume IV: Descriptive Epidemiology*. *IARC Scientific Publication No. 128*. Lyon, IARC, 1994.
25. 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.
26. Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J, eds. *Cancer Incidence in Five Continents. Volume VII. IARC Scientific Publication No. 143*. Lyon, IARC, 1997.
27. 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*. In press.
28. Brenner H, Hakulinen T. Up-to-date long term survival curves of patients with cancer by period analysis. *J Clin Oncol*. In press.