NOWOTWORY Journal of Oncology • 2004 • volume 54
Number 5 • 441–445

Invited review

The period analysis methodology in survival analysis: towards more up-to-date monitoring of long-term survival rates

Hermann Brenner

Long-term survival rates are increasing for certain forms of cancer in many countries. It would be highly desirable to detect such increases as timely as possible. A few years ago, period analysis has been introduced as a new methodology to derive more up-to-date estimates of long-term survival rates. After thorough empirical evaluation and development of user-friendly software, this methodology is now ready for implementation as a standard tool for up-to-date monitoring of cancer survival by cancer registries. This review gives a short outline of the new methodology, its statistical background, empirical evaluation, computational realization and applications.

Metoda analizy okresowej w analizie przeżycia – jak uaktualnić obserwacje dotyczące długotrwałych okresów przeżycia

W wielu krajach okresy przeżycia pacjentów z chorobą nowotworową ulegają wydłużeniu. Byłoby wskazane, aby takie zmiany wychwytywać tak wcześnie, jak tylko to możliwe. Przed kilkoma laty wprowadzono analizę okresową (period analysis) jako nową metodę, pozwalającą uzyskać bardziej aktualne dane dotyczące odsetka długotrwałych przeżyć. Po dokonaniu dokładnej oceny empirycznej i po opracowaniu „przyjaznego" oprogramowania, uwzględniającego tę metodę, można obecnie uznać, że jest ona gotowa do wprowadzenia jako standardowe narzędzie, umożliwiające uzyskanie aktualnych danych, dotyczących czasu przeżycia w oparciu o dane z rejestrów zachorowań na nowotwory. Niniejsza praca przedstawia krótki zarys tej nowatorskiej metody, z uwzględnieniem jej statystycznej podbudowy, oraz przedstawia wyniki jej oceny empirycznej, metody obliczeń i możliwe zastosowania.

Key words: neoplasms, prognosis, registries, survival
Słowa kluczowe: nowotwory, rokowanie, rejestry zachorowań, przeżycie

Introduction

Long-term survival rates are increasing for various forms of cancer in many countries. It would be highly desirable that cancer registries could detect such increases as timely as possible. However, traditional estimates of long-term survival, which are derived by cohort-based types of analysis [1-3], essentially reflected the survival expectations of patients diagnosed many years ago. A few years ago, a new methodology of survival analysis, denoted period analysis, has been introduced that enables more up-to-date estimates of long-term survival rates [4, 5]. This review gives a short outline of the new methodology, its statistical background, empirical evaluation, computational realization and applications.

Department of Epidemiology
German Centre for Research on Ageing
Heidelberg, Germany

The period principle

The principle of period analysis of survival is very simple. Period estimates of survival are derived for the most recent period for which data on cancer incidence and survival are available in the registry. For that purpose, all observations of patients included in the study are left truncated at the beginning of that period in addition to being right censored at its end. The principle and its differences from traditional survival analysis are illustrated in Figure 1.

Assume, a cancer registry has collected data from cancer patients diagnosed between the years 1980 and 2000, and that all patients have been followed up with respect to vital status until the end of the year 2000. Then, the most up-to-date estimate of 5-year-survival by traditional "cohort analysis" would pertain to patients diagnosed in 1995 and followed from 1995 to 2000. Analogously, the most up-to-date estimates of 10-, 15- and 20-year survival would pertain to patients diagnosed in 1990, 1985, and 1980, respectively, taking their survival experience from diagnosis to the year 2000 into account (black frames in

Year of Diag-nosis	Years of Diagnosis 1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
1980	1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12	12/13	13/14	14/15	15/16	16/17	17/18	18/19	19/20	20
1981		1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12	12/13	13/14	14/15	15/16	16/17	17/18	18/19	19/20
1982			1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12	12/13	13/14	14/15	15/16	16/17	17/18	18/19
1983				1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12	12/13	13/14	14/15	15/16	16/17	17/18
1984					1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12	12/13	13/14	14/15	15/16	16/17
1985						1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12	12/13	13/14	14/15	15/16
1986							1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12	12/13	13/14	14/15
1987								1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12	12/13	13/14
1988									1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12	12/13
1989										1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12
1990											1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11
1991												1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10
1992													1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9
1993														1	1/2	2/3	3/4	4/5	5/6	6/7	7/8
1994															1	1/2	2/3	3/4	4/5	5/6	6/7
1995																1	1/2	2/3	3/4	4/5	5/6
1996																	1	1/2	2/3	3/4	4/5
1997																		1	1/2	2/3	3/4
1998																			1	1/2	2/3
1999																				1	1/2
2000																					1

Figure 1. Principle of derivation of recent period estimates (grey frame) versus traditional cohort estimates (black frames) of long-term survival rates illustrated in the context of a cancer registry with incidence and follow-up data from 1980 to 2000. The numbers within the cells indicate the years of follow-up since diagnosis

Figure 1). This implies that the long-term survival estimates obtained from this database at the beginning of the 21st century might be quite out of date.

By contrast, the period approach would exclusively reflect the survival experience in the year 2000, regardless whether 5-, 10-, 15- or 20-year survival estimates are determined. Survival during the 1st year following diagnosis would be obtained from patients diagnosed in 1999 and 2000, conditional survival during the 2nd year following diagnosis would be obtained from patients diagnosed in 1999 and 1998, and so on, until conditional survival experience in the 20th year following diagnosis which would be obtained from patients diagnosed in 1980 and 1981 (grey frame in Figure 1). Cumulative survival rates are then obtained by multiplication of those conditional survival rates in the usual manner.

What the period principle achieves

Comprehensive empirical evaluation carried out within the last few years has shown that:

- long-term period survival estimates are much more up-to-date than traditional estimates of long-term survival estimates in case of improvements in survival over time,
- period survival estimates are identical to traditional survival estimates if prognosis does not change over time,
- long-term period survival estimates for a given period quite closely predict the survival rates observed many years later for patients diagnosed in that period,
- period analysis advances detection of trends in 5-, 10-, 15- and 20-year survival rates by almost 5, 10, 15 and 20 years, respectively, compared to traditional cohort survival analysis.

These patterns have been consistently shown for all of the most common forms of cancer among both adults [6-9] and children [10]. Although somewhat more up-to-date survival estimates can also be obtained by traditional survival analysis through the inclusion of more recently diagnosed patients who could not be under observation for the entire follow-up period of interest by the closing date of follow-up, even those "complete estimates" are typically much less up-to-date than the estimates that can be obtained by period analysis [7]. In case of very rapid increases of long-term survival rates over time, even the period estimates tend to remain slightly "too pessimistic", however, albeit much less so than the survival estimates obtained by traditional approaches.

As an example, figures 2 shows the most recent period estimates of 20-year relative survival curves for the year 2000 for women diagnosed with cancer in the United States below age 75 compared to the most recent 5-, 10-, 15- and 20- year relative survival curves obtained by cohort analysis (relative survival is obtained as the ratio of observed survival and expected survival in the absence of cancer, and it thereby reflects "net mortality" due to the cancer [3]). These estimates were derived from

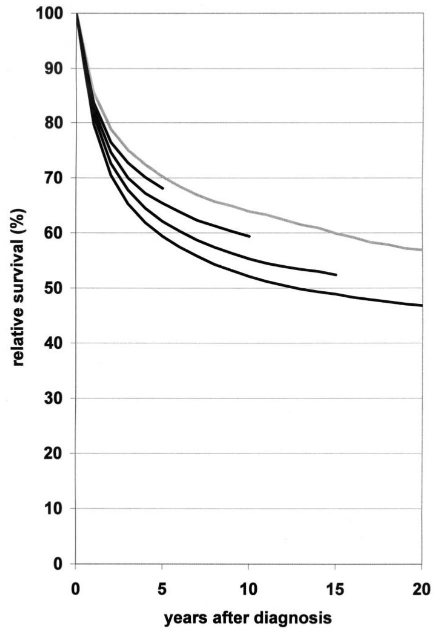

Figure 2. Recent 5-, 10-, 15- and 20-year relative survival curves for women diagnosed with any form of cancer below age 75 in the United States according to period analysis for the year 2000 (grey line) or according to cohort analysis (pertaining to patients diagnosed in 1995, 1990, 1985 and 1980, respectively) (black lines, from top to bottom)

the 1973-2000 public use database of the Surveillance, Epidemiology, and End Results (SEER) Program of the United States National Cancer Institute issued in April 2003 [11]. Relative survival was calculated according to the Hakulinen method [12]. The most recent period estimates of 5-, 10-, 15- and 20-year relative survival which can be obtained from a single survival curve, are 70, 64, 60 and 57%, respectively (grey curve in figure 2), whereas the most recent cohort estimates, which are obtained from different survival curves (black curves in figure 2), are between 2 and 10 percent units lower.

How to carry out period analysis

Period analysis can be conveniently carried out for both absolute (observed) and relative survival rates using publicly available SAS macros. Two macros are currently available, one called "period" [13] which can be used for calculation of absolute survival and for calculation of relative survival according to the so-called Ederer II method [14]. The other macro, called "periodh" [15], can be used for absolute survival and for relative survival calculated according to the Hakulinen method [12]. Both macros use a life-table approach with one-year intervals of follow-up. Numbers of deaths and persons at risk by follow-up year, as well as point estimates and standard errors (calculated according to Greenwood's formula [16]) of cumulative absolute, expected and relative survival rates can be provided. Apart from their use for period analysis, the macros can also be used for traditional (cohort-wise) life table analysis of survival.

The macros can be run with the SAS statistical software package version 8 or older. The macros and their documentation 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). The source code of the macros is open code under the conditions of the GNU-GPL license [17].

Recent Applications

In recent years, the period analysis approach has been adopted by an increasing number of countries to provide more up-to-date long-term cancer patient survival statistics. Period estimates of long-term cancer patient survival rates have now been reported for cancer patients in Finland [18], Estonia [19], Germany [20-24], the United Kingdom [25], and the United States [10, 26] and pertinent analyses are under way in a number of countries, including, among others, Denmark, Sweden, the Netherlands and Australia. Furthermore, period analysis is currently carried out in the ACCIS project, a collaborative project on childhood cancer survival in Europe [27]. Some cancer registries, such as the nationwide Finnish Cancer Registry and the nationwide German Childhood Cancer Registry, have already implemented period analysis as the standard tool for their annual reports [28, 29].

In all of the published applications, long-term survival estimates obtained by period analysis were considerably higher than long-term survival estimates obtained by traditional survival analysis for those forms of cancer for which major improvement in prognosis has been achieved over time. The advantages of period analysis compared to traditional cohort analysis have been found to be particularly large for childhood cancers, given the strong improvement in prognosis achieved by advancements in therapy for many forms of childhood cancer during the past decades. For example, recent period estimates of 15-year survival of children with acute lymphocytic leukemia and acute non-lymphocytic leukemia in Germany were 77% and 57%, respectively, compared to estimates of 69% and 36%, respectively, obtained by traditional cohort analysis [23].

Discussion

The period analysis methodology provides a useful tool to disclose progress in long-term survival of cancer patients, which has meanwhile been achieved for a number of cancers, and which is expected to continue in the future, in a more timely manner than traditional methods of survival analysis.

Like other methods of descriptive survival analysis, the period method does not by itself disclose the reasons for improvements in survival rates over time. Although such improvements are often due to advancements in therapy, as in the case of common forms of childhood cancers, they are sometimes also due to earlier diagnosis (e.g. in the context of introduction of new screening programs, such as screening for breast cancer by mammography), and they may even be due to overdiagnosis in certain instances (an issue that is of concern, for example, in the context of PSA screening for prostate cancer). However, this caveat has to be kept in mind in the interpretation of any survival statistics, and it is not specific to survival estimates obtained by period analysis.

This review focused on use of the period approach for descriptive survival analysis. The period approach may also be used in more complex analyses, such as multivariate modeling of survival rates [30]. Further developments may include the combination of period analysis with modeling of recent time trends or with other modeling approaches [e.g. 31, 32] to come up with even more up-to-date survival figures. However, the potential benefits of such extensions will have to be weighed against the increased complexity of methods and less forward interpretation of results.

Although the period approach to survival analysis may appear to be somewhat more difficult to digest and to communicate on first view than the traditional cohort-based survival analysis, the period approach is a straightforward and simple concept that has been successfully used for decades in other types of health statistics. In particular, the possibly most commonly cited and used parameter in health statistics, the life expectancy, is a classical period measure, as it is commonly derived from period life tables. Just like the life expectancy (derived from period life tables) quantifies the average length of life of a cohort of newborns on the basis of the (conditional) survival rates at various ages observed during some recent period (such as the most recent calendar year for which data are available), the period estimates of cancer patient survival quantify the survival expectations of newly diagnosed cancer patients on the basis of the most recently observed (conditional) survival rates in various years following diagnosis. There is no reason why period estimates should be less accepted for cancer survival rates than they have long been for the life expectancy.

In an era of steadily increasing survival rates of patients with many forms of cancer, provision of the most up-to-date survival estimates is not just an academic exercise. It may help to prevent patients, their relatives and clinicians from being unduly discouraged by outdated, often too pessimistic survival figures. Period analysis of survival, which should go along with efforts of high-quality and up-to-date cancer registration, may make an important contribution to this end.

Professor Hermann Brenner MD, PhD
Department of Epidemiology
German Centre for Research on Ageing
Bergheimer Str. 20
D-69115 Heidelberg
Germany
e-mail: Brenner@dzfa.uni-heidelberg.de

References

1. Cutler SJ, Ederer F. Maximum utilization of the life table method in analyzing survival. *J Chron Dis* 1958; 8: 699-712.
2. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958; 53: 457-81.
3. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: A statistical methodology. *Natl Cancer Inst Monogr* 1961; 6: 101-21.
4. Brenner H, Gefeller O. An alternative approach to monitoring cancer patient survival. *Cancer* 1996; 78: 2004-10.
5. Brenner H, Gefeller O. Deriving more up-to-date estimates of long term patient survival. *J Clin Epidemiol* 1997; 50: 211 -6.
6. Brenner H, Hakulinen T. Up-to-date survival curves of patients with cancer by period analysis. *J Clin Oncol* 2002; 20: 826-32.
7. Brenner H, Söderman B, Hakulinen T. Use of period analysis for providing more up-to-date estimates of long-term survival rates: empirical evaluation among 370 000 cancer patients in Finland. *Int J Epidemiol* 2002; 31: 456-62.
8. Brenner H, Hakulinen T. Advanced detection of time trends in long-term cancer patient survival: experience from 50 years of cancer registration in Finland. *Am J Epidemiol* 2002; 156: 566-77.
9. Brenner H, Hakulinen T. Very long-term survival rates of patients with cancer. *J Clin Oncol* 2002; 20: 4405-9.
10. Brenner H. Up-to-date survival curves of children with cancer by period analysis. *Br J Cancer* 2003; 88: 1693-7.
11. Surveillance, Epidemiology, and End Results (SEER) Program Public-Use Data (1973-2000). Bethesda (MD), National Cancer Institute, DCCPS, Cancer Surveillance Research Program, Cancer Statistics Branch, released April 2003, based on the November 2002 submission.
12. Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics* 1982; 38: 933-42.
13. Brenner H, Gefeller O, Hakulinen T. A computer program for period analysis of survival. *Eur J Cancer* 2002; 38: 690-5.
14. 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.
15. Brenner H, Hakulinen T, Gefeller O. Computational realization of period analysis for monitoring cancer patient survival. *Epidemiology* 2002; 13:611-2.
16. Greenwood M. A report on the natural duration of cancer. London: Ministry of Health, HMSO; 1926.
17. Free Software Foundation. The GNU General Public Licence. Available at: www.gnu.org/licenses/licenses.html. Accessed November 28, 2003.
18. Brenner H, Hakulinen T. Long-term cancer 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: 367-71.
19. Aareleid T, Brenner H. Trends in cancer patient survival in Estonia before and after the transition from a Soviet republic to an open market economy. *Int J Cancer* 2002; 102:45-50.
20. Brenner H, Stegmaier C, Ziegler H. Recent improvement in survival of breast cancer patients in Saarland, Germany. *Br J Cancer* 1998; 78: 694-7.
21. 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: 109-13.
22. Brenner H, Gefeller O, Stegmaier C et al. More up-to-date monitoring of long-term survival rates by cancer registries: an empirical example. *Methods Inf Med* 200l; 40: 248-52.
23. Brenner H, Kaatsch P, Burkhardt-Hammer T et al. Long-term survival of children with leukemia achieved by the end of the second millenium. *Cancer* 2001; 92: 1977-83.
24. Burkhardt-Hammer T, Spix C, Brenner H et al. Long-term survival of children with neuroblastoma prior to the neuroblastoma screening project in Germany. *Med Pediatr Oncol* 2002; 39: 156-62.
25. Smith LK, Lambert PC, Jones DR. Up-to-date estimates of long-term cancer survival in England and Wales. *Br J Cancer* 2003; 89: 74-6.
26. Brenner H. Long-term survival rates of cancer patients achieved by the end of the 20th century: a period analysis. *Lancet* 2002; 360: 1131-5.
27. ACCIS Automated Childhood Information System [Online]. Available from: http://www-dep.iarc.fr/accis/about.htm. Accessed November 25, 2003.
28. Finnish Cancer Registry. Cancer Incidence in Finland 1998 and 1999. Cancer Society of Finland Publication No. 63. Helsinki: Finnish Cancer Registry; 2002.
29. German Childhood Cancer Registry. Annual Report 2002. Mainz: German Childhood Cancer Registry; 2002.

30. Dickman PW, Sloggett A, Hills M et al. Regression models for relative survival. *Stat Med* (in press).
31. Mould RF, Lederman M, Tai P et al. Methodology to predict long-term cancer survival from short-term data using Tobacco Cancer Risk and Absolute Cancer Cure models. *Phys Med Biol* 2002; 47: 3893-924.
32. Mould RF. Prediction of long-term survival rates of cancer patients. *Lancet* 2003; 361: 262.

Paper received and accepted: 3 December 2003