ELSEVIER

European Journal of Cancer 40 (2004) 778–779

European Journal of Cancer

www.ejconline.com

Letter

Modification of SAS macros for a more efficient analysis of relative survival rates

Long-term survival rates are essential outcome measures of cancer care, and they are now reported by many cancer registries around the world (e.g. [1,2]). Some years ago period survival analysis was introduced by Brenner and Gefeller as a way to obtain more up-to-date long-term survival estimates [3]. Since then, the method has been extensively evaluated, and it has been used for deriving up-to-date estimates of long-term survival for patients diagnosed in recent years (e.g. [4]).

To facilitate the wider utilisation of period survival analysis, two SAS macros were developed by Brenner and colleagues [5,6]. The macros employ a life table approach to estimate cumulative absolute and relative survival rates. In their simplest form, the macros use a 2×100 array to store the age-specific population survival probabilities for males and females that are needed to estimate relative survival. As such, the same survival probabilities are used for every year of follow-up. This approach, although reasonable for most applications of period analysis where follow-up exclusively reflects survival during some short calendar periods (commonly 1–5 years), might be modified for cohort analysis where survival may be derived during a much longer calendar period (e.g. 20 years). To take changes in the expected survival probabilities of the underlying population over the years into account, Brenner and colleagues have suggested to use specific population survival probabilities for each calendar year included in the analysis by invoking pertinent macros containing this information at appropriate places within the program code [5].

Table 1
Results of the Performance Test (Period analysis of SEER-Data, Hakulinen-Method)

	Original Macro	Modified Macro	Difference
Hard disk space	218,120 pages	20,691 pages	−91%
Real Time	434.4 sec	88.0 sec	−80%
CPU-Time	81.4 sec	75.6 sec	−7%

SEER, Surveillance, Epidemiology and End Results. All analyses were performed with SAS© V8.02 on a AMD 1000 MHz PC with 256 Mbyte RAM, 40 GB EIDE hard drive (Western Digital WD400AB) and Windows 2000 professional© as the operating system.

However, referring to entire life tables increases the number of variables in the analysed data-set and may slow down calculations. Although this is of little concern for most practical applications, time and disk space requirements may become relevant or even prohibitive when very large data-sets are analysed. However, the performance of the macros may be substantially enhanced by minor modifications of the original macro code which we share with anyone that might find them useful. The modifications comprise the implementation of a temporary array to store the population survival probabilities, the compression of data files and the restriction of variables kept for analysis. In addition, we implemented a more clearly laid out tabular output that also facilitates the data transfer for further analyses including graphing data.

We tested the enhanced macro code in a re-analysis of recently published data from the Surveillance, Epidemiology and End Results (SEER) data base [4]. The period analysis for the year 1998 included patients diagnosed with any form of cancer between 1978 and 1998 in nine SEER registries ($n=1{,}730{,}564$). The new program code was much faster (−80%) and needed substantially less hard disc space (−91%) (see Table 1). We think that these modifications represent a significant improvement of the existing codes and should now enable the user to apply the period method to even extremely large-scale datasets. The updated macros are available free of charge at http://www.imbe.med.unierlangen.de/issan/SAS/period/period.htm. The SAS source code is open code under the conditions of the GNU-GPL license [7].

References

1. Berrino F, Capocaccia R, Estève J, *et al. Survival of Cancer Patients in Europe: the EUROCARE-2 Study. IARC Scientific Publication No. 151*. Lyon, IARC Press, 1999.
2. Sankaranarayanan R, Black RJ, Parkin DM. *Cancer Survival in Developing Countries. IARC Scientific Publication No. 145*. Lyon, IARC Press, 1999.
3. Brenner H, Gefeller O. An alternative approach to monitoring cancer patient survival. *Cancer* 1996, **78**, 2004–2010.
4. Brenner H. Long-term survival rates of cancer patients achieved by the end of the 20th century: a period analysis. *The Lancet* 2002, **360**, 1131–1135.
5. Brenner H, Gefeller O, Hakulinen T. A computer program for

0959-8049/$ - see front matter
doi:10.1016/j.ejca.2003.12.013

period analysis of cancer patient survival. *Eur J Cancer* 2002, **38**, 690–695.
6. Brenner H, Hakulinen T, Gefeller O. Computational realization of period analysis for monitoring cancer patient survival [letter]. *Epidemiology* 2002, **13**, 611–612.
7. Free Software Foundation. The GNU General Public License. Available at: www.gnu.org/licenses/licenses.html. Accessed August 22, 2003.

V. Arndt
Department of Epidemiology, German Centre for Research on Ageing, Bergheimer Strasse 20, D-69115 Heidelberg, Germany
E-mail address: arndt@dzfa.uni-heidelberg.de

M. Talbäck
Centre for Epidemiology, The National Board of Health and Welfare, Stockholm, Sweden
E-mail address: mats.talback@sos.se

O. Gefeller
Department of Medical Informatics, Biometry and Epidemiology, University of Erlangen-Nuremberg, Erlangen, Germany
E-mail address: olaf.gefeller@rzmail.uni-erlangen.de

T. Hakulinen
Finnish Cancer Registry; Department of Public Health, University of Helsinki, Helsinki, Finland
E-mail address: timo.hakulinen@cancer.fi

H. Brenner
Department of Epidemiology, German Centre for Research on Ageing, Heidelberg, Germany
E-mail address: brenner@dzfa.uni-heidelberg.de