

periodR – an R package to calculate long term cancer survival estimates using period analysis

Kongress Medizin und Gesellschaft 2007. Augsburg, 17.-21.09.2007

B Holleczek¹, A Gondos², H Brenner²

¹ Saarland Cancer Registry (Saarbrücken)

² German Cancer Research Center, Division for Clinical Epidemiology and Aging Research (Heidelberg)

periodR – an R package for survival estimation using period analysis

period analysis

- **long term survival** is key outcome parameter reported by cancer registries
- **period analysis**: reflects survival experience within a restricted recent calendar period („left truncation“)
- **method**: cumulative survival rates based on estimated conditional survival for subsequent time intervals after diagnosis (life table methodology)
- up-to-date prediction of results of traditional survival analysis (**early detection of changes in survival**)
- **empirically evaluated** and widespread use by cancer registries

periodR – an R package for survival estimation using period analysis

period analysis vs. traditional cohort based analysis

Diagnose	Nachbeobachtung												
	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
1992	1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10		
1993		1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10	
1994			1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10
1995				1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10
1996					1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9
1997						1	1/2	2/3	3/4	4/5	5/6	6/7	7/8
1998							1	1/2	2/3	3/4	4/5	5/6	6/7
1999								1	1/2	2/3	3/4	4/5	5/6
2000									1	1/2	2/3	3/4	4/5
2001										1	1/2	2/3	3/4
2002											1	1/2	2/3
2003												1	1/2
2004													1

traditional cohort based estimation of 10 year survival

period analysis based on observed survival between 2002-2004 (left truncation!)

mortality follow up available for end of period (2004)

periodR – an R package for survival estimation using period analysis



is a **language and environment for statistical computing and visualizing data**

- implements the S language
- available for **many OS** (Windows, Linux, MacOS)
- freely available under GNU public license
- **packages extend R** (provide additional **functions, datasets, documentation**)
- currently available version: R 2.5.1 (released 2007/06/28)
- september 2007: 1139 packages available from CRAN

periodR – an R package for survival estimation using period analysis

periodR

implements and extends **existing SAS macros** for period analysis

implementation/methodological details:

- calculation of **absolute and relative survival** based on conditional survival estimates for **1-year intervals** following diagnosis
- methods for relative survival: **Ederer II** and **Hakulinen**
- **standard errors** (Cutler-Ederer-method),
- **age-standardization** (uses individual weights)

periodR – an R package for survival estimation using period analysis

periodR

the package contains

- functions for **period analysis** (also cohort, complete or hybrid analyses possible)
- functions for data visualization (**plotting** survival curves)
- **datasets** (stomach cancer data, tables with age specific survival probabilities of males and females in Germany 1993-2002)
- integrated **documentation** and executable **examples**

periodR – an R package for survival estimation using period analysis

stomach cancer data

dataset of 2302 cancer cases diagnosed in Saarland between 1993 and 2002 (DCO cases already excluded)

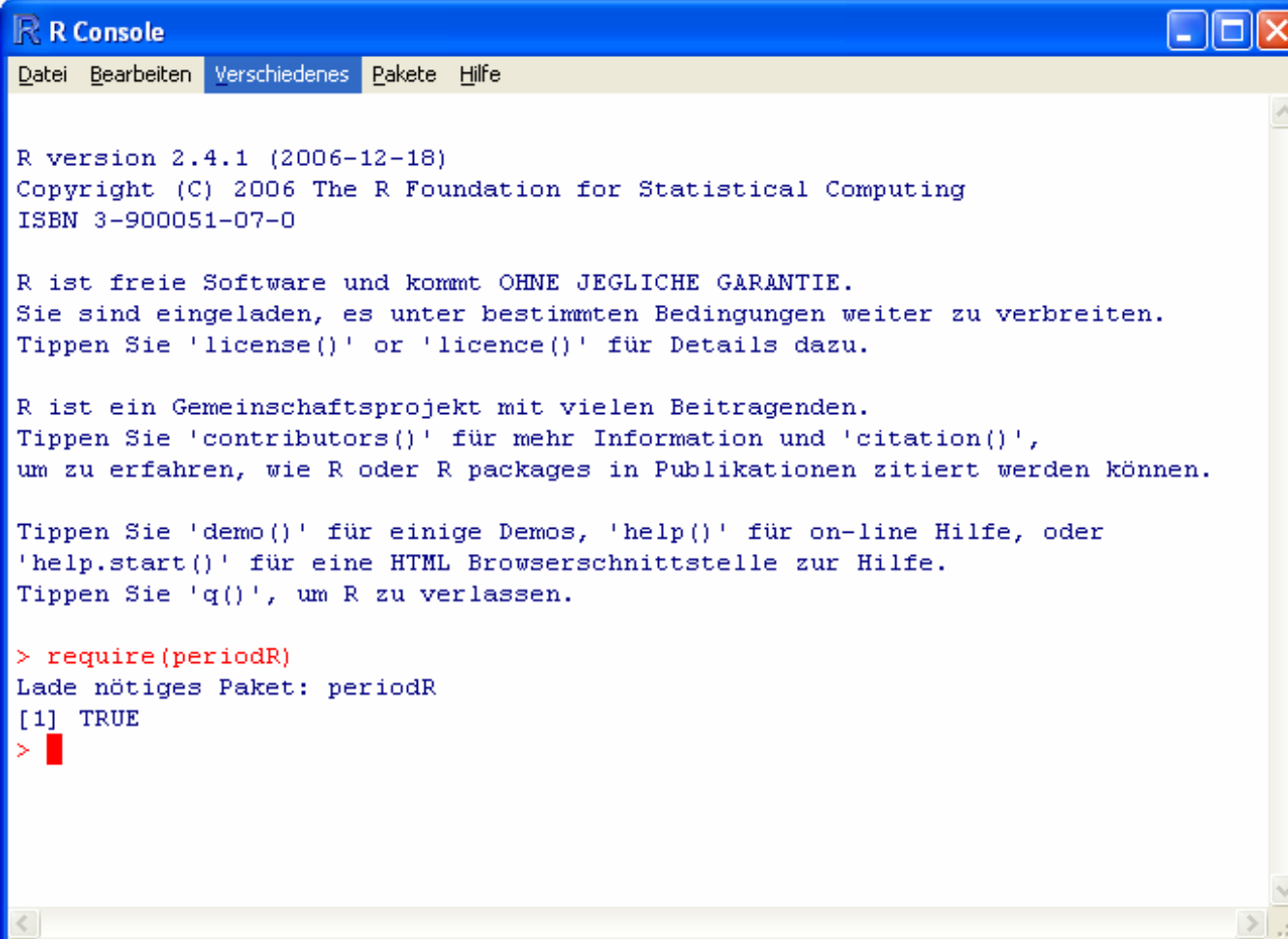
variables required for analysis

- sex
- age at diagnosis
- month and year of diagnosis
- month and year of end of follow up
- vital status at end of follow up

for relative survival needed: **lifetable data** of age and sex specific survival probabilities

periodR – an R package for survival estimation using period analysis

start of R session



```
R R Console
Datei Bearbeiten Verschiedenes Pakete Hilfe

R version 2.4.1 (2006-12-18)
Copyright (C) 2006 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R ist freie Software und kommt OHNE JEGLICHE GARANTIE.
Sie sind eingeladen, es unter bestimmten Bedingungen weiter zu verbreiten.
Tippen Sie 'license()' or 'licence()' für Details dazu.

R ist ein Gemeinschaftsprojekt mit vielen Beitragenden.
Tippen Sie 'contributors()' für mehr Information und 'citation()',
um zu erfahren, wie R oder R packages in Publikationen zitiert werden können.

Tippen Sie 'demo()' für einige Demos, 'help()' für on-line Hilfe, oder
'help.start()' für eine HTML Browserschnittstelle zur Hilfe.
Tippen Sie 'q()', um R zu verlassen.

> require(periodR)
Lade nötiges Paket: periodR
[1] TRUE
> █
```


periodR – an R package for survival estimation using period analysis

example analysis 1

```
> data(stomach)
> data(surv.probs.males)
> data(surv.probs.females)
>
> result <- period(data=stomach, k=10, surv.probs.males,
                   surv.probs.females, perbeg=2000, perend=2002,
                   method="hakulinen")
>
> print(result)
estimated survival using period analysis:
Call: period(data=stomach, k=10, surv.probs.males, surv.probs.females,
             perbeg=2000, perend=2002, method="hakulinen")

method used for relative survival: hakulinen
adjusted for age: no
observations included: 1164
```

periodR – an R package for survival estimation using period analysis

example analysis 2

estimates:

	year	abs.surv	err.abs	rel.surv	err.rel
1	1	0.5066	0.0197	0.5287	0.0206
2	2	0.3770	0.0191	0.4108	0.0208
3	3	0.3128	0.0183	0.3560	0.0209
4	4	0.2802	0.0178	0.3344	0.0212
5	5	0.2614	0.0175	0.3280	0.0219
6	6	0.2338	0.0169	0.3088	0.0224
7	7	0.2177	0.0165	0.3029	0.0230
8	8	0.2021	0.0161	0.2968	0.0237
9	9	0.1831	0.0161	0.2844	0.0251
10	10	0.1831	0.0161	0.3015	0.0266

year = follow up year

abs.surv = observed survival at end of follow up year

err.abs = standard error of observed survival

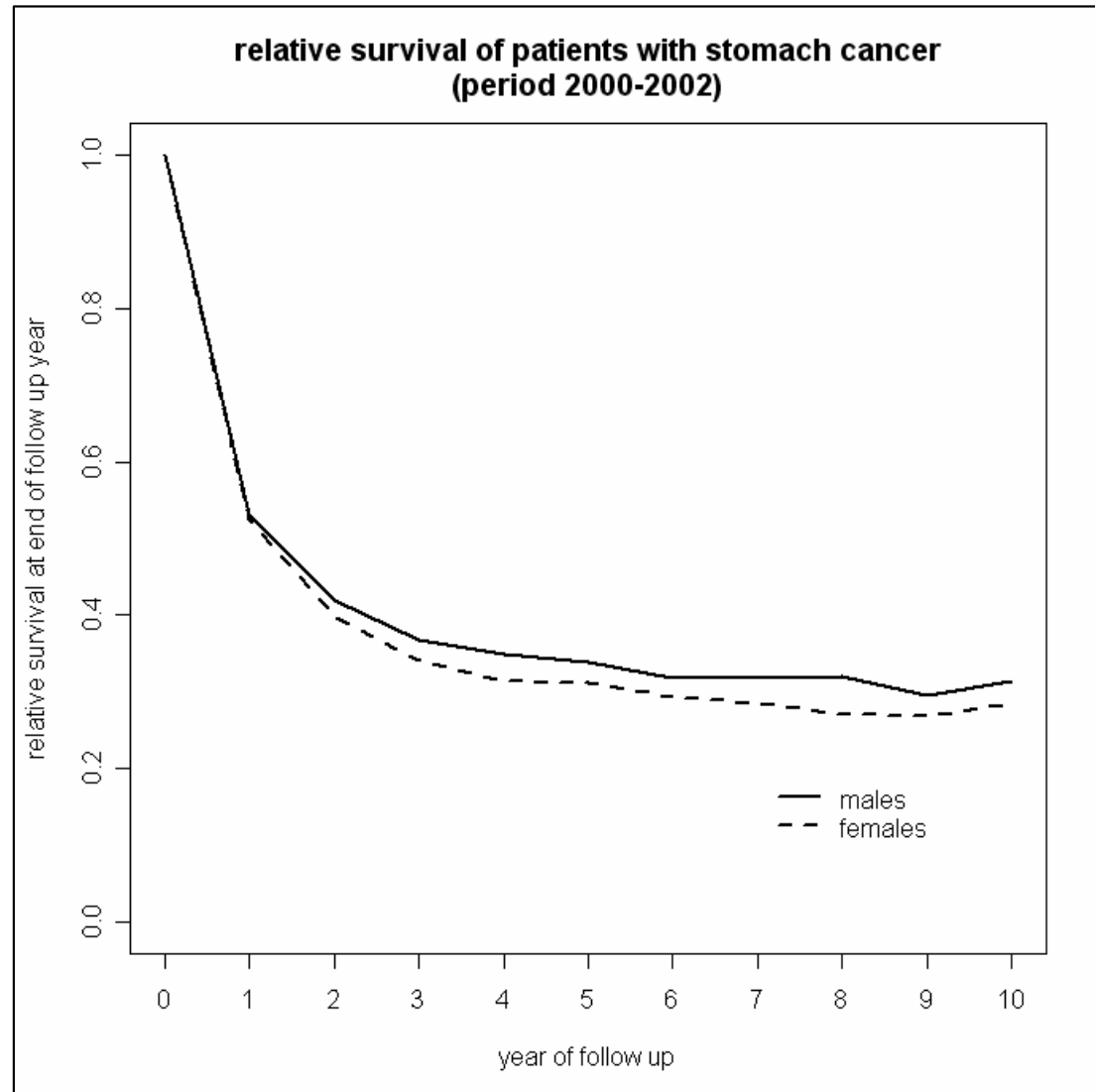
rel.surv = relative survival at end of follow up year

err.rel = standard error of relative survival

periodR – an R package for survival estimation using period analysis

data visualization

plotted survival curves
for both sexes for
period 2000-2002
(stomach cancer
data)



periodR – an R package for survival estimation using period analysis

integrated help

The image shows a screenshot of a Microsoft Internet Explorer browser window displaying the R website. The browser's address bar shows the path `C:\Programme\R\R-2.4.1\doc\html\index.html`. The main content of the browser window is the R website's index page, which features the text "Statistical Data Analysis" and the R logo. Below this, there are several links: "The R Language Definition", "R Installation and Administration", "R Internals", "Search Engine & Keywords", "Material", "Resources", and "Thanks".

Overlaid on the bottom-left corner of the browser window is a smaller window titled "R Help on 'period'". This window displays the help text for the 'period' package. The text is as follows:

```
Survival estimation using period analysis

Description:

  Calculation of survival estimates based on the period analysis
  approach

Usage:

  period(data, k, surv_males, surv_females, perbeg, perend, method=c("hakulinen", "edererII"), agedis$

Arguments:

data: dataframe consisting of the variables
  'diagage' (numeric vector: age in years at diagnosis),
  'dy' (numeric vector: year of diagnosis(e.g. 1995)),
  'dm' (numeric vector: month of diagnosis (e.g. 9 for
september)),
  'fy' (numeric vector: year of end of follow-up),
  'fm' (numeric vector: month of end of follow up),
  'vitstat' (numeric vector: vital status at end of follow-up
(1=alive, 2=dead)),
  'sex' (numeric vector: gender (1=male, 2=female))

k: integer variable: length of follow up in years (e.g. k=10 for
analysis of 10-year survival)

surv_males: dataframe of numeric vectors of length 100: each vector
contains conditional age-specific survival probabilities of
males during the periods specified by column names (e. g. if
dataframe contains three columns named "1990", "1994" and
"1995" then column 1 will be accessed for survival
```

periodR – an R package for survival estimation using period analysis

summary

periodR

- will be **freely available** under GNU public license
- intended to **enhance feasibility** and **application** of period analysis
- next step: **model based analysis**

further information/software

- <http://www.R-project.org> (R software, documentation, packages)
- <http://www.krebsregister.saarland.de/improve> (periodR)
- <http://www.imbe.med.uni-erlangen.de/issan/SAS/period> (SAS macros for absolute and relative survival rates)

periodR – an R package for survival estimation using period analysis

references

- Brenner H, Gefeller O. **An alternative approach to monitoring cancer patient survival.** Cancer 1996;78:2004-10.
- Brenner H, Hakulinen T. **Up-to-date long term survival estimates of patients with cancer by period analysis.** J Clin Oncol 2002;20:826-32
- Brenner H, Gefeller O, Hakulinen T. **Period analysis for up-to-date cancer survival data: theory, empirical evaluation, computational realization and applications.** Eur J Cancer 2004;40:326-35.