MEDINFO 2007
HISA Ltd Australia
© 2007 The authors. All rights reserved.

SAS Macro Program for Lead-time Bias and Length Bias Adjustment by Stochastic Modelling

Grace Hui-Min Wu[a, b], Tony Hsiu-Hsi Chen[b, c], Amy Ming-Fang Yen[b], Matti Hakama[a], Stephen D Walter[d], Anssi Auvinen[a]

[a]*Tampere School of Public Health, University of Tampere, Finland*
[b]*Institute of Preventive Medicine, National Taiwan University, Taiwan*
[c]*Division of Biostatistics, National Taiwan University, Taiwan*
[d]*Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada*

Abstract and objective

In evaluation of cancer screening, comparison of survival between cases detected by screening and cases detected clinically was often made as the effectiveness of screening. However, this comparison is not valid due to several biases including lead-time bias, length bias, etc.. Stochastic model depicting the process from normal, preclinical phase while the cases were detected by screening, via clinical phase to cancer death can be used to obtain the mean lead-time and real survival time for screen-detected cases which is free from lead-time bias and length bias. To facilitate the application of lead-time bias and length bias adjustment for survival by stochastic modelling, a SAS macro program was developed. The program is very flexible and enables the user to specify homogeneous and non-homogeneous (i.e. Weibull distribution, log-logistic, etc.) Markov models for each transition. The program was successfully applied to an example of prostate cancer screening.

Keywords:

Stochastic model, lead-time bias, length bias, screening, survival

Introduction

In evaluation of cancer screening, comparison of survival between cases detected by screening and cases detected clinically was often made as the effectiveness of screening [1-3]. However, this comparison is not valid due to several biases including lead-time bias, length bias, etc. [4-5]. For instance, it is possible that early detection of some prostate cancers may simply advance the date of diagnosis, i.e. lead-time gain, without prolonging life, namely lead-time bias. Moreover, screen-detected cases may include a higher proportion of slowly growing cancers with a long sojourn time, which results in length bias. Due to these sources of error, the magnitude of benefit achieved from early detection estimated by comparing the survival of screen-detected cases with that of clinically detected cases is elusive.

Stochastic model depicting the process from normal, preclinical phase while the cases were detected by screening, via clinical phase to cancer death, can be used to obtain the mean lead-time and real survival time for screen-detected cases which is free from lead-time bias and length bias [1]. Nevertheless, the complexity of computer computation of the Markov model may limit the use of this method. The aim of the current study is to develop a SAS macro program to facilitate the application of lead-time bias and length bias adjustment by stochastic modelling.

Methods

Model specification

The observed survival time of screen-detected cancers consists of lead-time gain and the real survival time in absence of screening, which would be from the onset of clinical disease (without lead-time), to cancer death. In comparison of the survival of screen-detected and clinically detected cases, survival from hypothetical detection in absence of screening should be used to eliminate the lead-time bias, i.e. ensure comparability of information. However, the time when the disease would have been identified clinically is unobservable for screen-detected cases. This can be solved by applying stochastic modelling.

Furthermore, to adjust for length bias, the probability of being diagnosed with an interval cancer (typically with a short sojourn time) together with the probability of being diagnosed at screen (with a long sojourn time) was incorporated to obtain an unbiased estimate of the sojourn time. Besides, in order to further adjust for the higher proportion of slowly growing cancers at the first screen, the probability of asymptomatic cancer or normal finding at screening has to be conditional on time of surfacing to clinical phase > time to screen.

Accordingly, we defined a five-state stochastic process (figure 1) with the combination of the disease natural history model, based on screen-detected cases (long sojourn time) and interval cancer (short sojourn time), as well as disease outcome model, based on screen-detected cases

only. The details of the method for length bias adjustment have been addressed elsewhere [6].

In the model (figure 1), $I(\tau)$, $\lambda_1(\tau)$ and $\lambda_2(\tau)$ represent the pre-clinical incidence rate, the instantaneous transition rate from preclinical phase to clinical phase, and hazard of dying from cancer, respectively. The competing cause of other deaths was taken into account by λ_3. Based on the estimates of $\lambda_2(\tau)$, the cumulative survival curve after correcting lead-time and length bias can be calculated as

$$S(t) = \exp\left(\int_0^t -\lambda_2(s)ds\right).$$

Maximum likelihood approach will be applied to estimate the transition parameters and the cumulative survival adjusted for lead-time and length bias can be also obtained.

SAS Macro ADJUST

The SAS macro ADJUST which has fourteen components has been developed. The description for each components are as follows.

(1) data, the SAS data set to be analyzed; (2) dist1, dist2, and dist3, the distribution for transition from normal to preclinical detectable phase, from preclinical detectable phase to clinical phase, and from clinical phase to die from the cancer, respectively, and the choices of the distribution for each transition include exponential distribution, weibull distribution, log-logistic distribution, and gamma distribution; (3) mode, the variable name for identify the different detection mode; (4) fn, the coding of variable "mode" for normal at first screen; (5) fsd, the coding of variable "mode" for screen-detected cancer at first screen; (6) sn, the coding of variable "mode" for normal at subsequent screen; (7) ssd, the coding of variable "mode" for screen-detected cancer at subsequent screen; (8) psfun, the coding of variable "mode" for normal during post-screening follow-up; (9) ic, the coding of variable "mode" for interval cancer; (10) ocd, the coding of variable "mode" for other cause of deaths of screen-detected cases; (11) event, the coding of variable "mode" for event, i.e. dying from the disease, of screen-detected cases; (12) censor, the coding of variable "mode" for censoring of screen-detected cases; (13) init, the initial values of each relevant transition parameter; (14) upcon and lowcon, the values of upper constraints and lower constraints on each relevant transition parameters.

Example: prostate cancer screening

The Finnish population-based prostate cancer screening randomized trial was started in 1996 [7]. During 1996 to 1999, 80,458 men aged 55-67 years and resident in the metropolitan areas of Helsinki and Tampere were identified from the Population Registry of Finland as the study population. Of them, 32,000 men were randomly allocated to the screening arm and were invited to PSA screening for prostate cancer. The men were recruited and invited for the first screening round in 1996-1999, with the second screening round carried out in 2000-2003, i.e. with a 4-year screening interval. The remaining 48,458 men were assigned to the control arm and were not contacted.

Men in the screening arm with a serum PSA concentration of 4.0 ng/mL or higher were referred to diagnostic examinations, consisting of digital rectal examination (DRE), transrectal ultrasound, and prostate biopsy. Men with a PSA concentration of 3.0-3.9 ng/mL were offered a secondary screening test, which was initially (1996-98) a DRE by an urologist and since 1999 determination of the proportion of free PSA [7]. Prostate cancers diagnosed in 1996-2002 in both the screening and control arms were identified and followed up until death or until end of 2002. Causes of death were obtained from death certificates.

Example of SAS Macro ADJUST

```
%ADJUST(DATA=c.prostate, DIST1=weibull,
DIST2=exponential, DIST3=weibull, MODE=type,
FN=1, FSD=2, SN=3, SSD=4, PSFUN=5, IC=6, OCD=7,
EVENT=8, CENSOR=9, INIT=0.0002 1.8 0.02 0.00001
2.3 0.02, LOWCON=0 0 0 0 0 0, UPCON=. . . . . .);
```

Results

The difference of survival between screen-detected cases and clinically detected cases has been heavily diminished after adjusting for lead-time bias and length bias. The unadjusted risk ratio of death for screen-detected cases against clinically detected cases was 0.17 (95% CI: 0.10-0.28). The risk ratio increased to 0.71 (95% CI: 0.51-1.00) after correcting for lead-time and length bias.

Conclusion

The program is very flexible and enables the user to specify homogeneous and non-homogeneous (i.e. Weibull distribution, log—logistic, etc.) Markov models for each transition. The program was successfully applied to an example of prostate cancer screening.

Reference

[1] Gilliland FD, Hunt WC, Key CR (1996) Improving survival for patients with prostate cancer diagnosed in the prostate-specific antigen era. Urology 48:67-71.

[2] Paquette EL, Sun L, Paquette LR, et al. (2002) Improved prostate cancer-specific survival and other disease parameters: impact of prostate-specific antigen testing. Urology 60:756-9.

[3] Quaglia A, Vercelli M, Puppo A, et al. (2003) Prostate cancer in Italy before and during the 'PSA era': survival trend and prognostic determinants. Eur J Cancer Prev 12:145-52.

[4] Eddy DM (1980) Screening for cancer: theory, analysis, and design. USA.

[5] Walter SD, Stitt LW (1987) Evaluating the survival of cancer cases detected by screening. Statistics in Medicine 6(8):885-900.
[6] Wu GHM, Chen THH, AMF Yen, Hakama M, Walter SD, Auvinen A. Survival of Screen- and Clinically-detected Prostate Cancer in a Population-based Screening Trial, Adjusted for Lead-time and Length Bias. (In reviewing in *Cancer Causes and Control*)
[7] Finne P, Stenman UH, Määttänen L (2003) The Finnish trial of prostate cancer screening: where are we now? BJU Int 92(Supl):22-26.

Address for correspondence

Professor Tony Hsiu-Hsi Chen, Institute of Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan. Room 521, No. 17, Hsu-Chow Rd, Taipei, Taiwan. Tel: +886-2-33228021; Fax: 886-2-23587707; E-mail: chenlin@ntu.edu.tw.

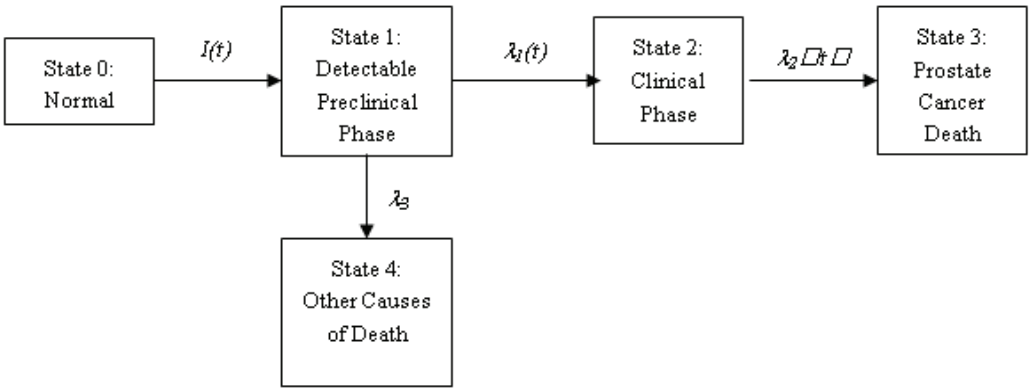

Figure 1 - Diagrams of a five-state Markov model combining the natural history and prognosis of prostate cancer from normal to death for length bias and lead-time adjustment