A Relative Survival Model for Clustered Responses

Kuss, O.[1], Blankenburg, T.[2], Haerting, J.[1]
[1] Institute of Medical Epidemiology, Biostatistics, and Informatics, University of Halle-Wittenberg, Halle (Saale)
[2] City Hospital "Martha-Maria" Dölau, Halle (Saale)
oliver.kuss@medizin.uni-halle.de

Introduction

Relative Survival is the ratio of the overall survival of a group of patients to the expected survival for a demographically similar group from a reference population, where this expected survival is derived from published age-, sex-, and calendar-time-specific mortality rates [1].

It is commonly used to estimate the effect of a particular disease when the true cause of death is not reliably known and is therefore the preferred analysis for survival experience in cancer registries, thus avoiding the problem of inaccurate or non-available death certificates [2].

The relative survival approach has several attractive features, for example, it allows for claiming cure (in a statistical, not a clinical sense) in the case where the relative survival in a group of patients equals the expected survival in the population. Moreover, comparisons between international registries are facilitated because by using relative survival the survival experience in different countries is adjusted for the respective underlying population.

Regression Models for Relative Survival

Generalizing the pure description of relative survival, regression models for relative survival have been proposed [2,3] to judge influence of prognostic or risk factors on relative survival. Dickman et al. [4] have shown that the relative survival model of Estève [2] can also be interpreted and fitted as a Generalized Linear Model with a Poisson response, an offset and a specific link function which is different for each observation:
Consider i=1,…, I observations, each one contributing j=1,…, J_i years of observations until death or censoring, δ_{ij} event indicator with δ_{ij} ~ Poisson(μ_{ij}), r_{ij} time at risk in the ij-th interval, e_{ij} expected numbers of death in the population.

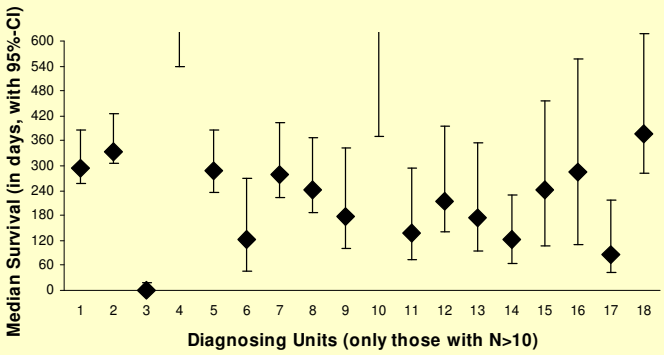
$$ln(\mu_{ij} - e_{ij}) = ln(r_{ij}) + x_i\beta$$

HALLUCA-Study

- (=Hallesche Lungen-Carcinom-Studie) Observational study which investigated provision of medical care of lung cancer patients in the region of Halle. All lung cancer patients in the study region were recorded from April 1996 to September 1999, follow-up was until September 2000.
- N=1696, 1349 death, Median survival: 285 days (9.3 months)
- Age- and sex-specific mortality rates were achieved from the Federal Statistical Office of the State of Saxony-Anhalt.

The Problem

Very heterogeneous survival experience in the 55 different diagnosing units in our study region (see figure below for the 18 largest units)



A Relative Survival Model for Clustered Responses

Idea: Generalize the Dickman GLM model to a mixed model with a unit-specific random effect u_h with u_h~N(0,σ^2). Additionally, allow for overdispersion by using a negative binomial distribution of the response [5].

$$ln(\mu_{hij} - e_{hij}) = ln(r_{hij}) + x_{hi}\beta + u_h$$

Parameter estimation by SAS PROC NLMIXED, model comparison by BIC (smaller is better).

Results

Estimate (SE)		Standard Estève	Random Effects (RE)	Negative Binomial (NB)	Description
Gender (ref.: male)		-0.158 (0.074)	-0.161 (0.076)	-0.168 (0.081)	80.6% male
Age (ref.: <65 years)		0.120 (0.058)	0.118 (0.060)	0.136 (0.064)	48.7% <65 years
Histological type (ref.: SCLC)	NSCLC	0.119 (0.069)	0.120 (0.071)	0.121 (0.077)	69.7% NSCLC
	Missing	-0.142 (0.116)	-0.143 (0.120)	-0.150 (0.125)	8.7% Missing
Performance state (ECOG)	3 – 4	0.692 (0.108)	0.714 (0.114)	0.832 (0.134)	7.3% 3 – 4
(ref.: 0 – 2)	Missing	0.136 (0.062)	0.145 (0.065)	0.151 (0.069)	43.9% Missing
Tumor stage	II	0.519 (0.183)	0.528 (0.187)	0.541 (0.194)	4.7% II
(ref.: I)	IIIa	0.672 (0.143)	0.683 (0.146)	0.707 (0.151)	11.5% IIIa
	IIIb	1.051 (0.132)	1.074 (0.136)	1.119 (0.142)	16.5% IIIb
	IV	1.452 (0.122)	1.485 (0.126)	1.550 (0.132)	36.6% IV
	Missing	0.553 (0.134)	0.575 (0.138)	0.602 (0.142)	19.8% Missing
σ^2		---	0.053 (0.037)	0.037 (0.037)	---
k (overdispersion)		---	---	0.208 (0.095)	---
BIC		6820.2	6807.5	6809.4	---

Discussion

Comparing the results from the different models, it can be seen that the parameter estimates for the covariates are different, but do not lead to different conclusions in subject matter terms.
As expected, the parameter estimates for the covariates are further away from the null and have slightly larger standard errors in the random effects models (RE & NB).
Comparing the standard relative survival model to both random effects models, we see a clear fall in the BIC values and thus the random effects model are preferred, however, the additional complexity of the NB model is not needed.

References

[1] Buckley JD. Additive and multiplicative models for relative survival rates. Biometrics 1984; 40:51-62.
[2] Estève J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: Elements for further discussion. Stat Med 1990; 9:529-538.
[3] Hakulinen T, Tenkanen L. Regression analysis of relative survival rates. Appl Stat 1987; 36:309-317.
[4] Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression Models for Relative Survival. Stat Med 2004; 23:51-64.
[5] Booth JG, Casella G, Friedl H, Hobert JP. Negative Binomial Loglinear Mixed Models. Stat Modelling 2003; 3:179-191.