
Relative Überlebenszeitanalyse von Lungenkrebspatienten im südlichen Sachsen-Anhalt

Oliver Kuß, Andre Gohr, Thomas Blankenburg, Johannes Haerting
Institut für Medizinische Epidemiologie, Biometrie und Informatik
Universität Halle-Wittenberg

Workshop der AG Statistische Methoden in der Epidemiologie,
14.3. - 15.3. 2002,
Institut für Epidemiologie und Sozialmedizin, Universität Münster

Programm

1. Einleitung

2. Relative Überlebenszeitanalyse

3. Regressionsmodelle für Relative Überlebenszeitanalyse

4. Ergebnisse für die HALLUCA-Studie

5. Software

6. Diskussion

7. Literatur

1. Einleitung

Die Überlebenszeitanalyse in Tumorregistern muss sich, im ist von der in klinischen Studien zu unterscheiden

Drei Möglichkeiten, um das Überleben in Tumorregistern zu beschreiben:

- Rohe/beobachtete Überlebenszeitanalyse
Problem: Nicht alle Patienten sind am relevanten Tumor verstorben
- Todesursachen-spezifische Überlebenszeitanalyse
Problem: Angaben auf den Totenscheinen müssen korrekt und eindeutig sein
- **Relative Überlebenszeitanalyse**
Inzwischen Quasi-Standard in Tumorregistern

2. Relative Überlebenszeitanalyse

Prinzip: Standardisierung des Überlebens der Registerpatienten an der Allgemeinbevölkerung durch Vergleich von beobachtetem und erwartetem Überleben

Interpretation:

- Überleben in einer hypothetischen Population, in der der Tumor die einzige Todesursache ist (Quasi-Todesursachen-spezifisches Überleben)
- Maß für die Übersterblichkeit in der Patientengruppe (unabhängig davon, ob der untersuchte Tumor die wahre Todesursache ist)

Eigenschaften:

- Information über Todesursachen nicht notwendig
- Zusätzliche Komplikation: Populationsdaten werden benötigt
- Annahme: Registerpatienten sind repräsentativ für die Population
- „Heilung“ (im statistischen Sinne) kann beschrieben werden
- Vergleich von verschiedenen Registern wird erleichtert

Schätzung von Relativen Survival-Raten:

Eigentliches Problem: Schätzen der erwarteten Überlebenszeiten

- 3 Methoden:**
- Ederer I** (1961): unverzerrt, aber Annahme von identischen Follow-up Zeiten
 - Ederer II** (1959): erlaubt heterogenes Follow-up, hängt aber von der beobachteten Mortalität ab
 - Hakulinen** (1982): korrigiert die Schwächen der Ederer-Methoden, benötigt aber Information über potentiell Follow-Up

Prinzipiell basieren aber alle drei Verfahren auf Berechnung von gewichteten Mitteln über die erwarteten Hazards der Patienten, es unterschieden sich nur die Art der Gewichte

3. Regressionsmodelle für Relative Überlebenszeitanalyse

Reines Beschreiben des relativen Überlebens reicht nicht aus, man möchte auch Faktoren, die das relative Überleben beeinflussen, identifizieren, deren Einfluß quantifizieren und statistisch prüfen.

(Stichwort: „Cox-Modell“ für relative Überlebenszeiten)

3 Modelle:

- Hakulinen/Tenkanen (1987)
- Estève (1990)
- Weller (1999)

Allen gemeinsam: Additive Hazard Regression

$$\lambda_{obs} = \lambda_{pop} + \lambda_{excess}$$

mit λ_{obs} : beobachteter Hazard,

λ_{pop} : aus der amtlichen Bevölkerungsstatistik und

$\lambda_{excess} = \exp(X\beta)$: Excess-Hazard, Funktion der Kovariablen

vgl. Cox-Modell:

$$\lambda_{obs} = \lambda_0 \exp(X\beta)$$

multiplikatives Modell, Baseline-Hazard λ_0 wird nicht geschätzt

Hakulinen/Tenkanen-Modell:

1. Gruppiere Patienten in K Strata, die durch Kreuzklassifikation aller Kovariablen definiert ist
2. Annahme: Im i -ten Zeitintervall im k -ten Stratum ist die Anzahl der Todesfälle binomialverteilt mit Sterbewahrscheinlichkeit $1-p_{ki}$
3. Einsetzen in Gleichung liefert:

$$\ln\left(-\ln\frac{p_{ki}}{p_{ki}^*}\right) = X\beta \quad p_{ki}^* = \text{erwartete Sterbewahrscheinlichkeit}$$

→ Generalisiertes lineares Modell mit Anzahl der Überlebenden im ki -ten Intervall/Stratum als Zielgröße, binomialem Fehler und komplementärer LogLog-Linkfunktion.

Allerdings: Linkfunktion für alle Strata und Intervalle unterschiedlich!!! (Software, die das erlaubt, wird benötigt)

4. Ergebnisse für die HALLUCA-Studie

HALLUCA (= **H**allesche **L**ungen-**C**arcinom-Studie): Versorgungsepidemiologische Studie in Zusammenarbeit mit dem klinischen Tumorregister

Standardisierte Erfassung aller Lungenkrebspatienten in der Region Halle zwischen 4/1996 und 9/1999 (Follow-Up bis 9/2000)

N=1549, 1252 Todesfälle, 297 Zensierungen, Medianes Überleben: 9.2 Monate

Faktoren im Modell: (alle kategorial, a priori festgelegt)

Geschlecht	Männlich	80,6 %				
Alter	<65 Jahre	48,7 %				
Histologischer Typ	SCLC	23,6 %				
Leistungszustand	0-2	50,7%	3-4	7,2%	Fehlend	42,1%
Tumorstadium	I	10,9%	II	4,9%	IIIa	12,1%
	IIIb	17,6%	IV	37,4%	Fehlend	17,1%

Populationsdaten vom Statistischen Landesamt Sachsen-Anhalt, 2 Intervalle

Ergebnisse I: Hakulinen-Modell/ Cox-Regression

		Hakulinen EHR[95%-KI]	Cox HR[95%-KI]
Geschlecht (Ref.: Männl.)		0.89 [0.77, 1.03]	0.88 [0.77, 1.02]
Alter (Ref.: <65 Jahre)		1.14 [1.01, 1.28]	1.19 [1.06, 1.33]
Hist. Typ (Ref.: SCLC)		0.89 [0.77, 1.01]	0.88 [0.77, 1.01]
Leistungszust. (Ref.: 0-2)	3-4	1.54 [1.22, 1.94]	1.79 [1.44, 2.22]
	Fehlend	1.15 [1.01, 1.31]	1.17 [1.04, 1.32]
Tumorstadium (Ref.: Stadium I)	II	1.76 [1.22, 2.54]	1.68 [1.20, 2.35]
	IIIa	1.98 [1.48, 2.65]	1.87 [1.44, 2.44]
	IIIb	2.77 [2.12, 3.63]	2.62 [2.05, 3.34]
	IV	4.18 [3.25, 5.36]	3.98 [3.12, 4.90]
	Fehlend	1.77 [1.34, 2.33]	1.70 [1.32, 2.18]

Ergebnisse II: Estève/Weller

		Estève EHR[95%-KI]	Weller EHR[95%-KI]
Geschlecht (Ref.: Männl.)		0.89 [0.77, 1.04]	0.88 [0.75, 1.04]
Alter (Ref.: <65 Jahre)		1.14 [1.01, 1.28]	1.17 [1.03, 1.33]
Hist. Typ (Ref.: SCLC)		0.88 [0.76, 1.02]	0.85 [0.74, 0.99]
Leistungszust. (Ref.: 0-2)	3-4	1.85 [1.48, 2.31]	1.78 [1.39, 2.28]
	Fehlend	1.19 [1.04, 1.35]	1.19 [1.03, 1.36]
Tumorstadium (Ref.: Stadium I)	II	1.81 [1.15, 2.83]	1.88 [1.26, 2.82]
	IIIa	2.05 [1.44, 2.91]	2.16 [1.57, 2.97]
	IIIb	2.90 [2.05, 4.10]	3.26 [2.43, 4.37]
	IV	4.38 [3.18, 6.03]	5.16 [3.94, 6.78]
	Fehlend	1.85 [1.32, 2.59]	1.86 [1.38, 2.53]

5. Software

- SAS-Makros (Therneau et al., 1994, modifiziert von Baade)
Problem: Nicht validiert, schwer durchschaubar, nutzt Approximationen
- SURV2/SURV3 (Voutilainen et al., 1998)
Entwickelt im Finnischen Krebsregister, berechnet „nur“ Überlebensraten, Regression mit SAS, GLIM oder S-Plus
- RELSURV (Hédelin, 1997)
Berechnet das Estève-Modell

6. Diskussion

- Keine großen Unterschiede zwischen Risikofaktoren für rohes und relatives Überleben gefunden
- Wahrscheinlicher Grund: Mittlere Überlebenszeit bei Lungenkrebs zu kurz, als dass Unterschiede sichtbar werden könnten
- Ergebnisse für Estève- und Weller-Modell sehr ähnlich
- Software existiert, aber deren Verwendung ist alles andere als eine Freude (Land in Sicht: P. Dickman implementiert alles in SAS)

7. Literatur

- Ederer F, Axtell LM, Cutler SJ. The relative survival rate:A statistical methodology.National Cancer Institute Monograph ,6:101 –121,1961.
- Ederer F, Heise H. Instructions to IBM 650 programmers in processing survival computations.Methodological note No.10,End Results Evaluation Section,National Cancer Institute,Bethesda MD,1959.
- Estève J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: Elements for further discussion. *Statistics in Medicine* , 9:529 –538, 1990.
- Hakulinen T.Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics* , 38:933 –942, 1982.
- Hakulinen T, Tenkanen L. Regression analysis of relative survival rates. *Applied Statistics* ,36:309 –317, 1987.
- Hédelin G. RELSURV 2.0 a program for relative survival analysis .Department of Epidemiology and Public Health, Faculty of Medicine, Louis Pasteur University, Strasbourg, France, 1997.
- Voutilainen ET, Dickman PW, Hakulinen T. SURV2: Relative Survival Analysis Program (software manual), Helsinki: Finnish Cancer Registry, 1998
- Therneau T, Sicks J, Bergstralh E, Offord J. Expected survival based on hazard rates. Rochester: Mayo Clinic, 1994.
- Weller EA, Feuer EJ, Frey CM, Wesley MN. Parametric relative survival regression using generalized linear models with application to Hodgkin's lymphoma, *Applied Statistics*, 48:79-89, 1999.