



INSTITUT FÜR STATISTIK
SONDERFORSCHUNGSBEREICH 386



Aydemir, Hornsteiner:

Analyse zeitveränderlicher Kovariablen und
rekurrenter Ereignisse am Beispiel einer Studie zur
prophylaktischen Behandlung von
Oesophagusvarizen

Sonderforschungsbereich 386, Paper 81 (1997)

Online unter: <http://epub.ub.uni-muenchen.de/>

Projektpartner



Analyse zeitveränderlicher Kovariablen und rekurrenter Ereignisse am Beispiel einer Studie zur prophylaktischen Behandlung von Oesophagusvarizen

Sibel Aydemir * und Ulrich Hornsteiner **

Abstract

To assess the effect of prophylactic sclerotherapy on variceal hemorrhage and survival of patients with liver cirrhosis and esophageal varices, a randomized study had been carried out. We analysed the data from different points of view.

The time-dependent Cox model and the linear counting process of Aalen are applied allowing for the time-dependent covariate “variceal bleeding” – that switches up to three times – in a multivariate analysis of the remaining life time. A model for the times to and between the recurrent events of bleeding including unobserved heterogeneity is estimated by a distribution-free and by a parametric method where the latter also admits time-dependent covariates such as repeated measurements of laboratory data.

We find that high age, high Child-Pugh score and especially the first occurrence of variceal bleeding have a statistically significant negative effect on survival whereas patients with fundic varices and/or alcoholic cirrhosis have a significant higher risk of bleeding. In both analyses, inclusion of time-dependent covariates does not change the estimation substantially. In particular, prophylactic sclerotherapy is not shown to reduce the risk of bleeding nor dying significantly.

*Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie, Ludwig-Maximilians-Universität München; SFB 386, Teilprojekt B3.

**Lehrstuhl für Statistik, Wirtschaftswissenschaftliche Fakultät, Universität Regensburg; SFB 386, Teilprojekt B1.

1 Einleitung

Am Beispiel einer Studie zur Untersuchung des Effekts der prophylaktischen Sklerosierung bei Patienten mit Oesophagusvarizen werden in der vorliegenden Arbeit Verfahren aus den Teilprojekten B1 und B3 diskutiert. Oesophagusvarizen sind Krampfadern in der Speiseröhre, die typischerweise als Folge einer Leberzirrhose entstehen. Eine besondere Gefahr geht von den Varizen im Falle ihrer Blutung aus, die häufig auch zum Tode führt. Fraglich ist, ob eine prophylaktische Sklerosierung, also Verödung der Varizen *vor* dem Auftreten einer Blutung, eine sinnvolle Therapie darstellt.

Ziel der in sieben Münchener Kliniken durchgeführten Studie war es also, den Einfluß einer derartigen Prophylaxe auf das Überleben der Patienten sowie auf die Zeit bis zum Auftreten von Varizenblutungen zu untersuchen. Untersuchungen hinsichtlich des Auftretens der *ersten* Blutung sowie univariate Survivalanalysen sind in Sauerbruch et al. (1989) publiziert. Blutungen der Oesophagusvarizen können innerhalb der Beobachtungszeit aber auch mehrmals auftreten, wobei der Zeitpunkt einer Blutung natürlich zufällig ist. Eine ausführliche Beschreibung der Oesophagusvarizen-Daten, die neben fixen potentiellen Prognosefaktoren auch verschiedene zeitveränderliche Einflußgrößen enthalten, erfolgt in Abschnitt 2.

Die Analyse der Überlebenszeit unter Berücksichtigung der zeitlichen Struktur der Prognosefaktoren, speziell der Varizenblutungen, wird in Abschnitt 3 dargestellt. Die Schätzungen werden mit dem in Teilprojekt B3 behandelten time-dependent Cox-Modell und dem auf zeitveränderliche Kovariablen adaptierten linearen Modell von Aalen vorgenommen. Von besonderem Interesse ist der Vergleich dieser zeitveränderlichen Schätzungen mit entsprechenden fixen Analysen, die die zeitliche Struktur der Varizenblutungen ignorieren.

Abschnitt 4 enthält Untersuchungen der Zeit bis zum Auftreten von Varizenblutungen, die mit den in Teilprojekt B1 bearbeiteten Methoden aus dem Gebiet der multivariaten Ereignisanalyse rekurrenter Ereignisse durchgeführt werden. Das verwendete Accelerated-Failure-Time-Modell beinhaltet die Möglichkeit unbeobachteter Heterogenität. Die Schätzung erfolgt zum einen durch ein verteilungsfreies und zum anderen durch ein parametrisches Verfahren. Auch hier wird der Schätzung anhand zeitkonstanter Kovariablen die Berücksichtigung einer zeitveränderlichen Einflußgröße gegenübergestellt.

2 Die Oesophagusvarizen-Daten

Es handelt sich hierbei um eine randomisierte klinische Studie, in die in den Jahren 1982 bis 1986 insgesamt 133 Patienten mit diagnostizierter Leberzirrhose und Oesophagusvarizen rekrutiert wurden. Die Patienten wurden zufällig in zwei Therapiegruppen eingeteilt, wobei eine Gruppe ($n=68$) prophylaktisch sklerosiert wurde, die Kontrollgruppe ($n=65$) hingegen keine prophylaktische Behandlung erhielt. Innerhalb der Beobachtungszeit verstarben insgesamt 55 Patienten (41.4 %), davon 32 Patienten aus der Kontrollgruppe und 23 therapierte Patienten.

Tabelle 1: *Univariate Verteilung der Prognosefaktoren*

Geschlecht		Therapie	
männlich	n= 88	Beobachtung (Kontrolle)	n= 65
weiblich	n= 45	Sklerosierung	n= 68
Diagnose		Fundusvarizen	
nicht-alkohol. Zirrhose	n= 46	keine	n=118
alkoholische Zirrhose	n= 87	vorhanden	n= 15
Pughpunkte		Alter	
5- 6 Pughpunkte	n= 48	≤ 49 Jahre	n= 43
7- 9 Pughpunkte	n= 59	50 -64 Jahre	n= 60
10-15 Pughpunkte	n= 26	≥ 65 Jahre	n= 30

Neben der Therapiegruppe und dem Auftreten von Blutungen wurden potentielle Prognosefaktoren wie Geschlecht, Alter, Art der Diagnose, Vorhandensein von Fundusvarizen, Anzahl der Pughpunkte, Hämoglobin- und Bilirubinwerte erhoben. Für Hämoglobin und Bilirubin liegen neben den Meßwerten zum Zeitpunkt der Aufnahme in die Studie zusätzlich auch Messungen aus follow-up Untersuchungen vor.

Unter den verschiedenen Diagnosearten, wie z.B. cryptogene oder post-hepatitische Zirrhose, ist die alkoholische Zirrhose mit 65.4 % am häufigsten vertreten. In den nachfolgenden Analysen wird auf eine strengere Differenzierung der Diagnoseart verzichtet und lediglich zwischen alkoholischer und nicht-alkoholischer Zirrhose unterschieden.

Fundusvarizen bezeichnen im Oesophagus vorkommende Varizen, die bis hin zum Magenfundus reichen.

Die Kovariable „Pughpunkte“ entspricht dem Child-Pugh-Score und gibt den Schweregrad der Lebererkrankung wieder. Die Child-Pugh-Klassifikation ist in drei Gruppen eingeteilt, wobei mit aufsteigender Klassifikation und somit zunehmender Anzahl an Pughpunkten auch der Grad der Erkrankung zunimmt.

Die Verteilung der kategorialen und der kategorisierten Einflußgrößen ist Tabelle 1 zu entnehmen.

Die Verteilungen der Hämoglobin- und Bilirubinwerte bei Aufnahme in die Studie sehen folgendermaßen aus: Der mittlere Hämoglobinwert beträgt 13.2 g/dl, der mediane Hämoglobinwert 13.4 g/dl. Für Bilirubin liegt der mittlere Baselinewert bei 3.1 mg/dl, der Median bei 1.9 mg/dl. Hämoglobin und Bilirubin wurden in mehreren Untersuchungen im Laufe der Studie (max. 10 follow-up Untersuchungen) bestimmt. Die Normalwerte betragen für Hämoglobin 14-18 g/dl, Hämoglobinwerte unter 10 g/dl gelten als pathologisch. Der Normalbereich für Bilirubin liegt bei maximal 0.2-1 mg/dl, Werte über 2 mg/dl werden als pathologisch erhöht betrachtet.

Blutungen der Oesophagusvarizen konnten innerhalb der Beobachtungszeit bei insgesamt 41 Patienten (30.8 %) festgestellt werden. Lediglich 9 Patienten bluteten dabei öfter als einmal (5 Patienten hatten insgesamt 2 Blutungen, 4 Patienten bluteten insgesamt dreimal). Während in der nicht-prophylaktisch behandelten Kontrollgruppe in 38.5 % der Fälle Blutungen registriert wurden, betrug der Anteil bei prophylaktisch

sklerosierten Patienten 23.5 %.

Auffällig in den vorliegenden Daten ist der hohe Anteil der Patienten mit alkoholischer Zirrhose in der jüngsten Altersgruppe ($p < 0.001$). Während bei den unter 50-jährigen Patienten in 88% der Fälle die Zirrhose alkoholischer Ursache ist, beträgt der Anteil der an alkoholischer Zirrhose erkrankten Patienten in der mittleren Altersgruppe nur 60 % und bei über 65-jährigen Patienten sogar nur noch 40%.

3 Die Analyse der Überlebenszeit

3.1 Univariate Analysen von potentiellen Prognosefaktoren

Ziel dieses Projekts war es, den Einfluß verschiedener potentieller Faktoren auf die Überlebenszeit zu untersuchen. Neben Alter, Geschlecht, Auftreten von Fundusvarizen, Anzahl der Pughpunkte, Diagnose und Therapie zählen zu den wichtigen Prognosefaktoren die teilweise auch mehrmals auftretenden Varizenblutungen. Die Zeitpunkte der einzelnen Blutungen sind dabei zufällig.

Der Einfluß der Prognosefaktoren wurde zunächst in univariaten Analysen untersucht. Statistisch signifikante Unterschiede bezüglich der Überlebenszeit resultierten zwischen den drei Altersstufen (Log-Rank, $p=.0002$), den Pughpunkte-Kategorien (Log-Rank, $p=.0053$) und Patienten mit bzw. ohne Fundusvarizen (Log-Rank, $p=.0001$). Dagegen ergab sich kein signifikanter Unterschied bezüglich der Überlebenszeit zwischen Patienten mit alkoholischer Zirrhose und Patienten, deren Krankheitsursache nicht alkoholisch bedingt war (Log-Rank, $p=.93$). Die Kaplan-Meier-Schätzungen zeigten ferner weder auffällige geschlechts- noch therapiespezifische Unterschiede. Das 25%-Quantil der Überlebenszeit (entspricht einer Überlebenswahrscheinlichkeit von 0.75) betrug bei nicht-therapierten Patienten zwar nur 300 Tage, bei sklerosierten Patienten immerhin 530 Tage, statistisch ist dieser Unterschied jedoch nicht signifikant (Log-Rank, $p=.0791$).

Ein bedeutender Prognosefaktor des Überlebens bei Patienten mit Leberzirrhose ist die Varizenblutung. Von 92 Patienten, die keine Blutung hatten, starben 29 Patienten (31.5%). Hingegen betrug die Eventrate bei Patienten mit Varizenblutungen das zweifache. Hier starben 26 (63.4%) der 41 Patienten, die innerhalb der Beobachtungszeit mindestens eine Blutung hatten. Der Unterschied bezüglich der Überlebenszeit zwischen Patienten mit und ohne Varizenblutungen ist ebenfalls auffällig (Log-Rank, $p=.0008$).

3.2 Multiple Analyse der Überlebenszeit

Multiple Analysen der Überlebenszeit wurden mit dem linearen Modell von Aalen und im Vergleich dazu mit multiplen Cox-Modellen, d.h. bei ausschließlich fixen Kovariablen mit dem Cox-proportional-hazards-Regressionsmodell und im Falle von zeitveränderlichen Kovariablen mit dem time-dependent Cox-Modell, durchgeführt. Ausführliche Beschreibungen der Modellansätze sind unter anderem in Aalen (1989, 1993), Aydemir et al. (1996a, 1996b) und Mau (1988) zu finden. Zur Schätzung des linearen Modells wurde das im Rahmen des Teilprojekts B3 erstellte SAS-Macro DynaSURV eingesetzt.

In den multiplen Analysen wurden zwei verschiedene Ansätze verfolgt: zum einen ein zeitveränderlicher Ansatz, der die zeitliche Struktur der Varizenblutungen berücksichtigt, und dem gegenüber ein fixer Ansatz, bei dem die Blutung als fixe Kovariable in die Analyse einbezogen wird. Im fixen Ansatz wird also lediglich unterschieden, ob der Patient eine Blutung hatte oder nicht, ohne Beachtung des Zeitpunktes der Blutung.

Bisher konnte in verschiedenen vorangegangenen Analysen, wie z.B. an den Daten der Stanford Heart Transplantation Studie (Aydemir et al., 1996a) gezeigt werden, daß bei Vorliegen von zeitveränderlichen Kovariablen die Berücksichtigung ihrer zeitlichen Struktur durchaus zu neuen Erkenntnissen führen kann und im Vergleich dazu die Nichtbeachtung der zeitlichen Kovariablenstruktur oftmals stark verzerrte Ergebnisse liefert.

Inwieweit in dem vorliegenden Datensatz eine Berücksichtigung der zeitlichen Struktur der Blutungen neue Erkenntnisse gegenüber einer fixen Analyse liefert, ist Gegenstand der anschließenden Untersuchungen und Diskussion.

3.2.1 Schätzung mit dem linearen Modell

Das lineare Intensitätsmodell zum Zeitpunkt t ist definiert als

$$\lambda(t) = Y(t)\alpha(t),$$

wobei $\alpha(t)$ den Vektor der unbekanntes Regressionsfunktionen bezeichnet. Der Ansatz der Analyse besteht in der Betrachtung kumulierter Regressionsfunktionen

$$A(t) = \int_0^t \alpha(s)ds,$$

die über einen Zählprozeßansatz

$$A^*(t) = \sum_{T_k \leq t} X(T_k)I_k$$

geschätzt werden, mit $X(t) = (Y(t)'Y(t))^{-1}Y(t)'$ der generalisierten Inversen der Datenmatrix $Y(t)$. Die Steigung der gegen die Zeit geplotteten geschätzten kumulierten Regressionsfunktionen zu einem bestimmten Zeitpunkt gibt den Einfluß der Kovariablen in dem betrachteten Moment wieder.

Die Schätzung der kumulierten Regressionsfunktionen und der entsprechende Test auf Signifikanz kumulierter Regressionsfunktionen bezogen auf die gesamte Beobachtungsdauer sind im Detail in Aydemir et al. (1996a) dargestellt.

In den nachfolgenden multiplen Analysen wird zunächst die zeitliche Struktur der Kovariable Blutung mitberücksichtigt. Verglichen mit anderen zeitveränderlichen Kovariablen, wie z.B. Transplantation, kommt bei der Kovariable Varizenblutungen erschwerend hinzu, daß zum einen bis zu drei Blutungen, d.h. drei Switches auftreten können, und zudem noch berücksichtigt werden sollte, daß aus medizinischer Sicht mehrfache Blutungen nicht gleichbedeutend sind, d.h. das Risiko der ersten Blutung nicht gleichzusetzen ist mit dem Risiko einer zweiten Blutung.

Tabelle 2: *Asymptotisch standardnormalverteilte Teststatistiken des linearen Aalen-Modells bei zeitveränderlichem Einbezug der Blutungen*

	1. Blutung	1. und 2. Blutung	alle 3 Blutungen
Geschlecht	-1.92	-1.86	-1.85
Alter	3.82	3.50	3.54
Therapie	-1.38	-1.63	-1.67
Diagnose	-0.47	-0.40	-0.41
Fundusvarizen	0.56	0.95	0.94
Pughpunkte	2.66	2.67	2.70
Blutung	4.12	3.65	3.47

Es wird das lineare Modell geschätzt mit den fixen Kovariablen Geschlecht, Alter, Therapie, Diagnose, Pughpunkte und Fundusvarizen, und der zeitveränderlichen Kovariable Blutungen, wobei im ersten Ansatz lediglich die erste Blutung, in weiteren Ansätzen die zweite und schließlich auch die dritte Blutung betrachtet werden.

Die Schätzungen der sechs fixen Kovariablen werden durch diese unterschiedliche Behandlung der Varizenblutungen kaum beeinflusst (vgl. Tabelle 2). Unabhängig davon, ob nur die erste, die ersten beiden oder alle drei Blutungen berücksichtigt werden, resultieren für die fixen Kovariablen sehr ähnliche Ergebnisse. Höheres Alter und eine höhere Anzahl von Pughpunkten haben auf einem Signifikanzniveau von 5% stets einen positiv auffälligen Einfluß auf den Hazard und wirken sich somit ungünstig auf das Überleben aus. Die Kovariablen Geschlecht und Therapie besitzen keinen statistisch auffälligen Einfluß auf den Hazard, jedoch ist der Trend erkennbar, daß eine prophylaktische Sklerosierung das Überleben begünstigt, und Frauen im Vergleich zu Männern eine günstigere Prognose besitzen. Die geschätzten kumulierten Regressionsfunktionen der Kovariablen Alter, Pughpunkte, Geschlecht und Therapie sind in Abbildung 1 dargestellt.

Für die Kovariablen Fundusvarizen und Diagnose ist in der multiplen Analyse kein Einfluß auf den Hazard zu erkennen. Varizenblutungen besitzen, ganz gleich, ob nur die erste Blutung oder auch spätere Blutungen in die Analyse eingehen, stets einen signifikant ungünstigen Einfluß auf das Patientenüberleben. Im Gegensatz zu den fixen Kovariablen wird die Effektschätzung der zeitveränderlichen Kovariable Blutung allerdings durch die Anzahl der modellierten Switches deutlich beeinflusst. Wird lediglich die erste Varizenblutung der Patienten mit ihrer zeitlichen Struktur berücksichtigt, so besitzt die kumulierte Regressionsfunktion einen steileren Verlauf als bei Berücksichtigung auch der nachfolgenden Blutungen. Vor allem in den ersten 200 Tagen nach der Randomisierung ist in diesem Fall ein stärkerer Einfluß auf den Hazard erkennbar (vgl. Abbildung 2). Das von Varizenblutungen ausgehende Risiko erscheint somit am größten, falls nur die erste Blutung in die Analyse eingeht und weitere Blutungen vernachlässigt werden. Diese „Risikominderung“ durch eine zweite und dritte Blutung liegt zum Teil daran, daß viele Patienten bereits kurz nach der ersten Blutung versterben und folglich nur noch „gesündere“ Patienten eine weitere Blutung erleben können. Betrachtet man beispielsweise ausschließlich Patienten, die nur eine Blutung hatten (n=32), so beträgt die nach der Blutung noch verbleibende mediane „Rest“-Beobachtungs- bzw. „Rest“-Überlebenszeit 21 Tage. Patienten, die öfter geblutet haben (n=9), befinden

Abbildung 1: Schätzungen der kumulierten Regressionsfunktionen mit dem linearen Modell

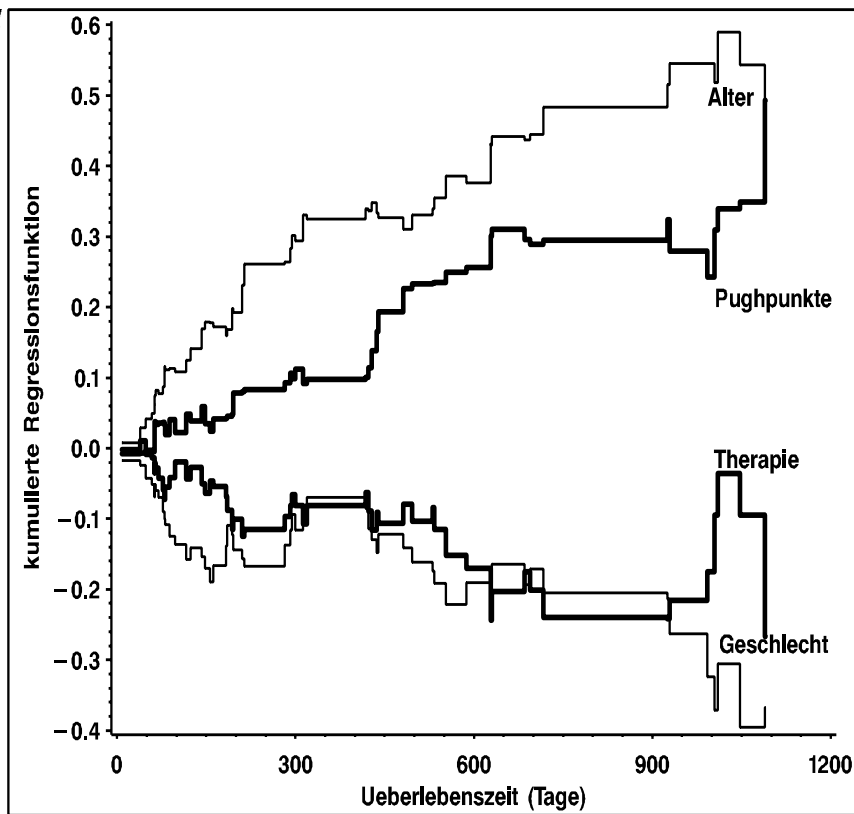


Abbildung 2: Schätzungen der kumulierten Regressionsfunktion der Kovariable Blutung für unterschiedliche Modellierungen

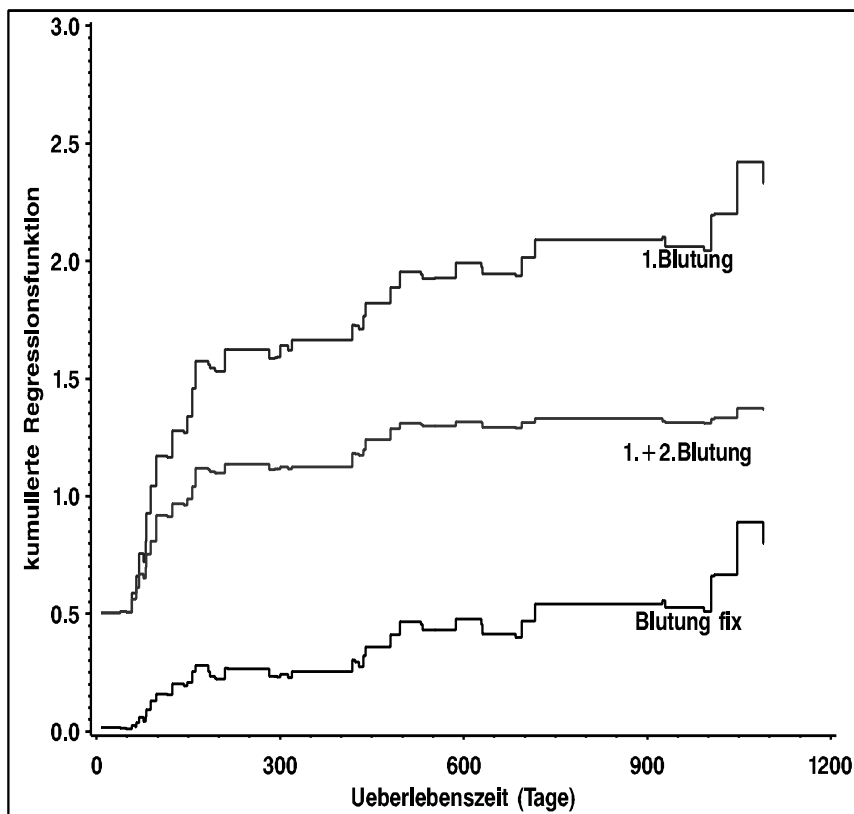


Tabelle 3: *Asymptotisch standardnormalverteilte Teststatistiken des linearen Aalen-Modells bei fixem Einbezug der Kovariable Blutung*

	Blutung fix
Geschlecht	-1.82
Alter	3.52
Therapie	-1.43
Diagnose	-0.37
Fundusvarizen	1.52
Pughpunkte	2.85
Blutung	3.02

sich insgesamt in einem vergleichsweise besseren gesundheitlichen Zustand, zumal sie die vorangegangene Blutung überlebt haben. Die nach der zweiten Blutung noch verbleibende mediane „Restzeit“ beträgt in diesem Kollektiv 91 Tage.

Eine weitere Modellierungsmöglichkeit bestünde darin, jede einzelne Varizenblutung durch Dummy-Variablen darzustellen. Dies hätte den Vorteil, daß für jede Blutung separate Effektschätzungen resultieren würden. Aufgrund des seltenen Vorkommens zweiter bzw. dritter Blutungen sind die vorliegenden Daten allerdings für diesen Ansatz ungeeignet.

Die kumulierten Regressionsfunktionen der Kovariable Blutung besitzen in den Analysen, in denen die ersten beiden oder alle drei Blutungen berücksichtigt werden, einen nahezu identischen Verlauf. Ein Grund hierfür ist sicherlich die zu geringe Anzahl der dritten Blutungen (insgesamt 4 Fälle).

In Tabelle 2 sind die asymptotisch standardnormalverteilten Teststatistiken der multiplen Analysen mit dem linearen Modell für verschiedene Switchzustände der zeitveränderlichen Kovariable Blutung aufgeführt.

Bei fixer Modellierung der Kovariable Blutung weichen die kumulierten Regressionsfunktionen der Kovariablen Geschlecht, Alter, Therapie, Diagnose, Pughpunkte und Fundusvarizen kaum von den entsprechenden Schätzungen aus den zeitveränderlichen Analysen ab. Wie bereits beim zeitveränderlichen Einbezug der Blutungen besitzen auch im fixen Ansatz höheres Alter und eine größere Anzahl von Pughpunkten einen signifikant positiven Einfluß auf den Hazard (vgl. Tabelle 3).

Varizenblutungen wirken sich bei fixer Modellierung ebenfalls auffällig ungünstig auf das Überleben aus. Allerdings verläuft ihre kumulierte Regressionsfunktion flacher als in den zeitveränderlichen Ansätzen. Verglichen mit den vorangegangenen Analysen, die die zeitliche Struktur der Blutungen berücksichtigen, resultiert bei fixem Einbezug der Kovariable Blutung folglich ein schwächerer Effekt. Daß Effektschätzungen bei Vernachlässigung der zeitlichen Kovariablenstruktur schwächer ausfallen, läßt sich häufig beobachten. Ursachen eines derartigen Effektschwundes werden in Altman & De Stavola (1994) diskutiert.

Abbildung 2 zeigt die geschätzten kumulierten Regressionsfunktionen der Kovariable Blutung für die verschiedenen Modellansätze. Die kumulierte Regressionsfunktion der Kovariable Blutung bei Berücksichtigung aller drei Varizenblutungen wurde aus

Gründen der Übersichtlichkeit nicht abgebildet, da sie sich nahezu überlappt mit der zeitveränderlichen Schätzung, in die die ersten beiden Blutungen eingehen (Abbildung 2, mittlere Funktion).

3.2.2 Schätzung mit dem Cox-Modell

Alternativ zum nicht-parametrischen linearen Modell von Aalen wird eine Analyse der Überlebenszeit mit dem semiparametrischen Cox-Regressionsmodell durchgeführt. Für den Fall, daß die Varizenblutungen als fixe Kovariable in die Analyse eingehen, wird das Cox-proportional-hazards-Regressionsmodell

$$\lambda_i(t) = \lambda_0(t) \exp\left[\sum_{j=1}^p \beta_j x_{ij}\right],$$

bei einem zeitveränderlichen Einbezug der Kovariable Blutung das sogenannte time-dependent Cox-Modell

$$\lambda_i(t) = \lambda_0(t) \exp\left[\sum_{j=1}^p \beta_j x_{ij}(t)\right]$$

angewendet. Dabei bezeichnet $\lambda_i(t)$ die Hazardfunktion des i -ten Individuums zum Zeitpunkt t und $\lambda_0(t)$ die unspezifizierte Baselinehazardfunktion.

Im Gegensatz zu den dynamischen Regressionsfunktionen $\alpha(t)$ des linearen Modells werden die aus dem Cox-Modell geschätzten Kovariableneffekte über den gesamten Beobachtungszeitraum hinweg als konstant angenommen. Der Regressionskoeffizient $\beta_j, j = 1, \dots, p$ stellt somit den Effekt der j -ten Kovariable zum Baselinezeitpunkt und zu allen späteren Zeitpunkten dar. Gegenüber dem klassischen Cox-Modell entfällt im time-dependent Ansatz jedoch – trotz der Restriktion auf konstante Kovariableneffekte – die Annahme der Proportionalität der Hazards. Detaillierte Beschreibungen der Cox-Modelle finden sich beispielsweise in Altman & De Stavola (1994) und Aydemir et al. (1996b).

Die Ergebnisse des Cox-Modells und des linearen Modells von Aalen stimmen gut überein. Auch in den multiplen Cox-Analysen besitzen Alter, Pughpunkte und Varizenblutungen einen signifikanten Einfluß auf den Hazard. Abhängig davon, ob eine fixe oder zeitveränderliche Analyse durchgeführt wird und wieviele Switches bezüglich der Kovariable Blutung berücksichtigt werden, schwanken die relativen Risiken für Alter zwischen 2.5 und 2.8 und für Pughpunkte zwischen 2.0 und 2.2. Größere Unterschiede bezüglich des relativen Risikos ergeben sich für die Varizenblutungen. Im fixen Ansatz besitzen Patienten mit Varizenblutungen ein 3.2-faches relatives Risiko zu versterben, verglichen mit Patienten, die keine Blutung hatten. Im zeitveränderlichen Modell, das lediglich die erste Blutung berücksichtigt, erhöht sich das relative Risiko auf das 8.4-fache. Hier zeigt sich also auch der stärkere Einfluß der Varizenblutungen bei zeitveränderlicher Modellierung gegenüber dem fixen Ansatz. Für die zeitveränderlichen Cox-Analysen, die die ersten beiden oder alle drei Blutungen berücksichtigen, resultieren – wie schon im linearen Aalen-Modell – sehr ähnliche Effektschätzungen. Für die ersten beiden Blutungen liefert das time-dependent Cox-Modell ein relatives Risiko von 4.4, bei Einbezug aller drei Blutungen ergibt sich ein relatives Risiko von 4.7.

Tabelle 4: *Ergebnisse der multiplen Cox-Analysen*

	Blutung fix	1. Blutung	1.+2. Blutung	alle 3 Blutungen
Geschlecht				
mle (std)	−0.54 (.31)	−0.65 (.32)	−0.75 (.33)	−0.73 (.33)
tst (p)	2.99 (.0839)	4.30 (.0382)	5.14 (.0234)	4.88 (.0272)
Alter				
mle (std)	0.93 (.22)	0.98 (.21)	0.98 (.21)	1.02 (.22)
tst (p)	18.42 (.0001)	21.95 (.0001)	20.91 (.0001)	22.01 (.0001)
Therapie				
mle (std)	−0.21 (.29)	−0.07 (.30)	−0.09 (.30)	−0.10 (.30)
tst (p)	0.51 (.4766)	0.06 (.8079)	0.10 (.7522)	0.11 (.7352)
Diagnose				
mle (std)	−0.23 (.35)	−0.30 (.35)	−0.39 (.36)	−0.42 (.36)
tst (p)	0.43 (.5108)	0.74 (.3908)	1.19 (.2763)	1.38 (.2396)
Fundusvarizen				
mle (std)	0.66 (.38)	0.40 (.41)	0.10 (.45)	0.02 (.45)
tst (p)	2.95 (.0860)	0.95 (.3309)	0.05 (.8277)	0.01 (.9615)
Pughpunkte				
mle (std)	0.77 (.22)	0.79 (.21)	0.70 (.21)	0.71 (.21)
tst (p)	12.36 (.0004)	13.56 (.0002)	11.29 (.0008)	11.77 (.0006)
Blutung				
mle (std)	1.16 (.33)	2.13 (.32)	1.48 (.25)	1.54 (.23)
tst (p)	12.67 (.0004)	43.51 (.0001)	34.06 (.0001)	43.21 (.0001)

Im Unterschied zum linearen Modell besitzt allerdings im time-dependent Cox-Modell, also nur in den zeitveränderlichen Analysen, auch Geschlecht einen signifikanten Einfluß auf den Hazard. Bei zeitveränderlicher Modellierung der Varizenblutungen ergibt sich demnach für Frauen ein auffällig geringeres Risiko zu versterben. Dieser signifikante Einfluß des Geschlechts zeigte sich weder in Analysen mit dem linearen Modell, noch im Cox-proportional-hazards-Modell bei fixer Modellierung der Blutung. Die Ergebnisse der multiplen Cox-Analysen sind in Tabelle 4 zusammengefaßt. Für unterschiedliche Betrachtungen der Kovariable Blutung sind ML-Schätzungen (mle) der Kovariableneffekte mit ihren Standardabweichungen (std), Wald-Teststatistiken (tst) und die entsprechenden p-Werte (p) aufgeführt.

4 Die Analyse der Verweildauern bis zur Blutung

4.1 Daten und Problemstellung

Da jede Oesophagusvarizenblutung mit ihrem Kalenderdatum dokumentiert ist, eignet sich der Datensatz auch für eine Analyse der Verweildauern bis zum teilweise pro Patient mehrmals wiederkehrenden Ereignis „Blutung“. Wir wollen also untersuchen, inwieweit man durch die im Datensatz gegebenen Größen (Alter, Geschlecht, Diagnose,

Vorhandensein von Fundusvarizen, Pughpunkte und Bilirubinwert ¹) die Häufigkeit der Blutungen vorhersagen kann und insbesondere, ob das Risiko von Blutungen durch die prophylaktische Sklerosierung verringert werden kann. Dazu benötigen wir Methoden aus dem Gebiet der multivariaten Ereignisanalyse rekurrenter Ereignisse.

Im folgenden sei n der Laufindex für die N Patienten, k der Laufindex für die K_n (teilweise) beobachteten Episoden des n -ten Patienten. Die Länge einer Episode T_{nk} ist dabei (exakt oder mindestens) die Zeit in Tagen

- von der Aufnahme des Patienten in die Studie bis zu seiner ersten Blutung,
- von einer Blutung bis zur nächsten,
- von der letzten dokumentierten Blutung bis zum Ende der Beobachtung des Patienten oder
- von der Aufnahme des Patienten in die Studie bis zum Ende der Beobachtung des Patienten, wenn der Patient keine Blutung hat.

In den beiden letzten Fällen liegt eine Rechtszensurierung der Episode vor und zwar unabhängig davon, ob der Patient stirbt, vorzeitig aus der Studie ausscheidet oder das Ende der festgelegten Beobachtungszeit erreicht ist. Bei den 13 Patienten, die bis zu einem Tag nach einer Blutung gestorben sind, wird als letzte Episode nur diejenige *bis* zur letzten Blutung berücksichtigt, weil davon ausgegangen werden kann, daß in diesen Fällen der Tod die unmittelbare Folge der Blutung war, eine zensierte Episode von der Länge eines Tages nur sehr geringe Information beinhalten würde und damit den ohnehin sehr hohen Anteil zensierter Episoden im Datensatz nur unnötig erhöhen würde.

Es stehen insgesamt 174 Episoden zur Auswertung zur Verfügung, von denen 120 (69 %) zensiert sind. Die durchschnittliche beobachtete Episodenlänge (incl. zensierter Werte) beträgt 505 Tage.

Alle Schätzalgorithmen dieses Abschnitts wurden in SAS/IML, Version 6 (SAS Institute Inc., 1989, 1990) programmiert.

4.2 Das Modell

Wir verwenden ein Accelerated-Failure-Time-Modell, bei dem der Logarithmus jeder Verweildauer direkt vom linearen Prädiktor aus dem Kovariablenvektor $x_{nk} = (1, x_{nk1}, \dots, x_{nkP})'$, den Regressionsparametern $\beta = (\beta_0, \beta_1, \dots, \beta_P)'$ und einer Störgröße abhängt (Fahrmeir, Hamerle und Tutz, 1996, S.310).

Dabei nehmen wir an, daß die Störgrößen zweier verschiedener Patienten unabhängig voneinander verteilt sind, solche eines Patienten jedoch miteinander korreliert sind. Dies wird im Modell erreicht durch eine Trennung der Störgröße in eine individuen-spezifische Heterogenitätskomponente α_n und einen individuen- und episodenspezifischen Teil ε_{nk} . Die Interpretation der Heterogenitätskomponente ist dabei meist die, daß sie Kovariablen repräsentiert, die bei einem Patienten konstant über die Zeit sind, aber nicht

¹Der Haemoglobinwert ist als Einflußgröße für Blutungen ungeeignet, da er sich eher als *Folge* einer Blutung verringert.

erhoben bzw. nicht in das Modell aufgenommen werden konnten. Das vollständige Modell lautet somit

$$y_{nk} = \ln T_{nk} = x'_{nk}\beta + \sigma_\alpha \alpha_n + \sigma_\varepsilon \varepsilon_{nk},$$

wobei wir die Standardabweichungen σ_α und σ_ε aus den Komponenten α_n und ε_{nk} herausnehmen, um deutlich zu machen, daß es sich dabei um weitere (außer den Regressionsparametern) unbekannte Parameter des Modells handelt.

Als Kovariablen sind in dieser Modelldarstellung sowohl solche zugelassen, die sich pro Patient im Zeitablauf nicht ändern, als auch solche, die in jeder Episode des Patienten einen anderen – innerhalb dieser Episode wiederum als konstant angenommenen – Wert annehmen. Die Kovariablen dieser Untersuchung sind fast alle solche, die sich pro Patient im Zeitablauf nicht ändern (Geschlecht und Art der Diagnose) oder von denen jedenfalls nur der Baseline-Zustand bekannt ist und dieser deshalb für die Auswertungen als konstant angenommen wird (Vorhandensein von Fundusvarizen und Pughpunkte). Auch die Variable Therapie wird über die Beobachtungszeit konstant gehalten ungeachtet der Tatsache, daß jeder Patient nach seiner ersten Blutung sklerosiert wurde (Sauerbruch et al., 1988). Dadurch wird hier nicht der Effekt einer Sklerosierung *nach* der ersten Blutung untersucht, was als Therapie zur Verhinderung weiterer Blutungen anerkannt ist, sondern der Effekt der Zugehörigkeit zur Therapiegruppe bzw. zur Kontrollgruppe und damit derjenige der *prophylaktischen* Sklerosierung auf das Auftreten der ersten und evtl. auch weiterer Blutungen.

Kovariablen, die sich während einer Episode ändern können, im vorliegenden Fall die Bilirubinwerte, werden in Abschnitt 4.5 im Modell berücksichtigt. Zunächst wird jeder Episode derjenige Bilirubinwert zugeordnet, der hinsichtlich des Episodenbeginns in der zeitlich am nächsten liegenden Verlaufsuntersuchung gemessen wurde.

Zum Modell gehört außerdem, daß wir bei zensierten Daten nicht die gesamte Verweildauer T_{nk} beobachten, sondern

$$z_{nk} = \min(T_{nk}, c_{nk}),$$

wobei der Zensierungsindikator

$$\delta_{nk} = \begin{cases} 1 & \text{falls } T_{nk} \leq c_{nk} \\ 0 & \text{falls } T_{nk} > c_{nk} \end{cases}$$

die Information enthält, ob es sich um eine zensierte Episode handelt oder nicht. Die Zensierungszeit c_{nk} hängt in diesem Fall über

$$c_{nk} = C_n - \sum_{l=1}^{k-1} T_{nl}$$

von der Gesamtbeobachtungsdauer C_n des Patienten ab, von der wiederum angenommen wird, daß sie nicht von den Verweildauern abhängt.

4.3 Die kombinierte GEE/Buckley-James-Schätzung

Erwünscht ist ein Schätzer für die Regressionsparameter, der sowohl die Information zensierter Episoden als auch die Abhängigkeitsstruktur der Verweildauern eines Patienten möglichst gut ausnutzt und dabei mit möglichst wenigen Verteilungsannahmen

auskommt.

In Hornsteiner und Hamerle (1996) wurde eine Methode mit solchen Eigenschaften vorgestellt. In Simulationsstudien hat sich gezeigt, daß dieser Schätzer gerade bei eher kleinen Stichproben von Untersuchungseinheiten mit aufeinanderfolgenden Episoden verhältnismäßig gute Ergebnisse liefert. Dies ist insbesondere der Fall bei Kovariablen, die über die Zeit hinweg konstant sind (dies trifft in dieser Untersuchung für alle Kovariablen außer dem Bilirubinwert zu) und bei nicht allzu starker Heterogenitätskomponente (was bei dieser Untersuchung ebenfalls zutrifft, wie sich zeigen wird). Dieses Verfahren benötigt im Gegensatz zu einer ML-Schätzmethode (siehe Abschnitt 4.4) keine Annahmen über die Verteilungen der Fehlerkomponenten und ist robust gegenüber einer falschen Spezifikation der Abhängigkeitsstruktur. Allerdings muß man bei einem sehr hohen Anteil zensierter Episoden (wie es bei den vorliegenden Daten der Fall ist) mit einer Überschätzung der Streuung der geschätzten Regressionsparameter und als Folge mit nicht entdeckten Effekten rechnen.

Grundlage der Methode bilden die *generalized estimating equations* (Liang und Zeger, 1986)

$$\sum_{n=1}^N X_n' V_n^{-1} (y_n^* - X_n \hat{\beta}) = 0,$$

wobei X_n die Zeilen x'_{nk} , $k = 1, \dots, K_n$ enthält, und $y_n^* = (y_{n1}^*, \dots, y_{nK_n}^*)'$ aus den *Pseudovariablen* (Buckley und James, 1979)

$$y_{nk}^* = \delta_{nk} \ln z_{nk} + (1 - \delta_{nk}) \hat{E}(y_{nk} \mid y_{nk} > \ln z_{nk})$$

gebildet wird. Somit ist für unzensierte Episoden $y_{nk}^* = y_{nk}$, während bei zensierten Daten der bedingte Erwartungswert der logarithmierten Verweildauer mit Hilfe des nicht-parametrischen Produkt-Limit-Schätzers (Kaplan und Meier, 1958) eingesetzt wird.

Ferner ist $V_n = R_{K_n}(\gamma)/\phi$ mit der „Arbeits-Korrelationsmatrix“ $R_{K_n}(\gamma)$ und dem Vektor γ , der die Korrelationsstruktur wiedergibt, und $1/\phi := v$ wird der Einfachheit halber als konstant angenommen. Als Korrelationsstrukturen eignen sich die Äquikorrelationsannahme und die Unabhängigkeitsannahme.

Die Schätzung erfolgt über einen dreistufigen, iterativen Algorithmus. Begonnen wird mit der Initialisierung $\hat{\beta}^{(0)} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{y}$, wobei \tilde{X} und \tilde{y} nur aus den unzensierten Episoden gebildet werden. Es folgen in der u -ten Iteration der Reihe nach die Aktualisierung der Buckley-James-Pseudovariablen, die Momentenschätzung von γ und v und schließlich das modifizierte Fisher-Scoring für β ,

$$\hat{\beta}^{(u)} = \hat{\beta}^{(u-1)} + \left(\sum_{n=1}^N X_n' \tilde{V}_n^{-1}(\hat{\beta}^{(u-1)}) \frac{\partial y_n^*}{\partial \beta} \right)^{-1} \left(\sum_{n=1}^N X_n' \tilde{V}_n^{-1}(\hat{\beta}^{(u-1)}) (y_n^* - X_n' \hat{\beta}^{(u-1)}) \right),$$

$u = 1, 2, \dots$, bis Konvergenzkriterien erfüllt sind. Der Algorithmus wird auch dann erfolgreich abgebrochen, wenn er nicht gegen einen Wert konvergiert, sondern sich in einer Schleife von mehreren Werten stabilisiert hat (vgl. hierzu auch Currie, 1996). Dies passiert umso häufiger und die Schleifen sind umso länger, je höher der Anteil der zensierten Episoden in den Daten liegt. Als Schätzergebnisse werden dann die Mittelwerte der relevanten Schleifenwerte berechnet (Miller, 1981, S. 152).

Tabelle 5: *Ergebnisse der GEE/Buckley-James-Schätzung mit Fundusvarizen, Diagnose und Therapie*

$\hat{\beta}$ [$\hat{\sigma}(\hat{\beta})$]	Konstante	8.34 [1.19]
	Fundusvarizen	-2.07 [0.819]
	Diagnose	-1.27 [1.34]
	Therapie	0.0304 [1.16]
\hat{v}		3.35
p-Werte	Fundusvarizen	0.013
	Diagnose	0.348
	Therapie	0.979

Als Schätzung für die asymptotische Kovarianzmatrix der Parameterschätzungen und für die Standardabweichungen wird

$$\widehat{Cov}(\hat{\beta}) = \left(\sum_{n=1}^N X_n' \tilde{V}_n^{-1} \frac{\partial y_n^*}{\partial \beta} \right)^{-1} \left(\sum_{n=1}^N X_n' \tilde{V}_n^{-1} \widehat{Cov}(y_n^*) \tilde{V}_n^{-1} X_n \right) \left(\sum_{n=1}^N X_n' \tilde{V}_n^{-1} \frac{\partial y_n^*}{\partial \beta} \right)^{-1}$$

und

$$\hat{\sigma}(\hat{\beta}_p) = \sqrt{(\text{diag}(\widehat{Cov}(\hat{\beta})))_p}$$

mit $\widehat{Cov}(y_n^*) = (y_n^* - x_n' \beta) (y_n^* - x_n' \beta)'$ verwendet.

Bei der Analyse der Oesophagusvarizen-Daten stößt man relativ schnell an die Grenzen der GEE/Buckley-James-Schätzmethode. Sie eignet sich zunächst zur groben Abschätzung von Richtung und Stärke der Effekte, indem man jeweils nur eine Variable in das Modell aufnimmt. Je mehr Variablen allerdings aufgenommen werden, desto länger sind die Schleifen, die der Schätzalgorithmus zyklisch durchläuft. Bereits bei drei Kovariablen (Fundusvarizen, Diagnose und Therapie) und Unabhängigkeitsannahme werden 900 Iterationen und drei Minuten Rechenzeit benötigt, um die Schleifenbildung zu entdecken. Der Grund dafür liegt in der sehr hohen Zensierungsquote bei relativ kleinem Stichprobenumfang. Zudem sind die geschätzten Standardabweichungen ziemlich hoch, wobei unbekannt ist, wie hoch der Anteil einer Überschätzung dieser Standardabweichungen liegt.

Ein solches Schätzergebnis ist in Tabelle 5 dokumentiert. Die Regressionsparameter für Fundusvarizen und Diagnose werden auf einen negativen Wert geschätzt, d.h. Patienten mit Fundusvarizen und solche mit alkoholischer Zirrhose haben kürzere Verweildauern bis zur Blutung als solche ohne Fundusvarizen und/oder anderer Diagnose. Allerdings ist nur der Effekt für die Fundusvarizen signifikant. Der positiv geschätzte Effekt der prophylaktischen Sklerosierung ist verschwindend gering.

4.4 Die Schätzung durch Maximierung der marginalen Likelihood

Setzt man im Vergleich zu 4.3 wesentlich restriktivere Annahmen bezüglich der Verteilungen der Fehlerkomponenten und ihrer Abhängigkeitsstruktur, so läßt sich ein Maximum-Likelihood-Schätzer für das gegebene Modell ableiten, der darauf basiert, über die unbekannte Heterogenitätskomponente hinweg zu integrieren und die – bezüglich der Heterogenität – marginale Likelihood zu approximieren. Dieses Verfahren wurde z.B. von Spieß und Hamerle (1995) im Kontext von Probitmodellen für abhängige diskrete endogene Variablen praktiziert. Dabei ist zu erwähnen, daß das Setzen solcher Annahmen nicht bedeuten muß, daß der resultierende Schätzer im Fall einer Fehlspezifikation unbrauchbare Ergebnisse liefert, vgl. hierzu Hornsteiner, Hamerle und Michels (1997). Die Annahmen lauten:

- Die α_n sind unabhängig und identisch standardnormalverteilt.
- Die ε_{nk} sind unabhängig und identisch verteilt mit beliebiger, bekannter Verteilung.
- α_n und ε_{nk} sind unabhängig.

Als mögliche Verteilungen für ε_{nk} werden im folgenden die drei in der Verweildauermanalyse gängigen verwendet, das sind Normalverteilung, logistische Verteilung und Extremwertverteilung, die zum Lognormal- (N), log-logistischen (L) bzw. Weibull-Modell (W) führen (Fahrmeir, Hamerle und Tutz, 1996, S. 310 ff.). Diese drei Modellversionen wurden auch im Schätzprogramm implementiert.

Die Log-Likelihood ist somit von der Gestalt

$$l(\beta, \sigma_\alpha, \sigma_\varepsilon \mid X_1, \dots, X_N, z_{11}, \dots, z_{N, K_N}, \delta_{11}, \dots, \delta_{N, K_N}) = \text{const} + \sum_{n=1}^N \ln \int_{-\infty}^{\infty} \exp \left(\sum_{k=1}^{K_n} (\delta_{nk} \cdot \ln f_{nk}(z_{nk}) + (1 - \delta_{nk}) \cdot \ln S_{nk}(z_{nk})) \right) \cdot \varphi(\alpha) d\alpha,$$

wobei die Dichtefunktion f_{nk} und die Survivorfunktion S_{nk} je nach Modellspezifikation folgende Formen annehmen:

(N)

$$f_{nk}(z_{nk}) = \frac{1}{\sigma_\varepsilon z_{nk}} \varphi \left(\frac{\ln z_{nk} - x'_{nk} \beta - \sigma_\alpha \alpha}{\sigma_\varepsilon} \right),$$

$$S_{nk}(z_{nk}) = 1 - \Phi \left(\frac{\ln z_{nk} - x'_{nk} \beta - \sigma_\alpha \alpha}{\sigma_\varepsilon} \right);$$

(L)

$$f_{nk}(z_{nk}) = \frac{1}{\sigma_\varepsilon} \exp \left(\frac{-x'_{nk} \beta - \sigma_\alpha \alpha}{\sigma_\varepsilon} \right) \cdot z_{nk}^{\frac{1}{\sigma_\varepsilon} - 1} \cdot \left(1 + \exp \left(\frac{-x'_{nk} \beta - \sigma_\alpha \alpha}{\sigma_\varepsilon} \right) \cdot z_{nk}^{\frac{1}{\sigma_\varepsilon}} \right)^{-2},$$

$$S_{nk}(z_{nk}) = \left(1 + \exp \left(\frac{-x'_{nk} \beta - \sigma_\alpha \alpha}{\sigma_\varepsilon} \right) \cdot z_{nk}^{\frac{1}{\sigma_\varepsilon}} \right)^{-1};$$

(W)

$$f_{nk}(z_{nk}) = \frac{1}{\sigma_\varepsilon} \exp\left(\frac{-x'_{nk}\beta - \sigma_\alpha\alpha}{\sigma_\varepsilon} - \exp\left(\frac{-x'_{nk}\beta - \sigma_\alpha\alpha}{\sigma_\varepsilon}\right) \cdot z_{nk}^{\frac{1}{\sigma_\varepsilon}}\right) \cdot z_{nk}^{\frac{1}{\sigma_\varepsilon}-1},$$
$$S_{nk}(z_{nk}) = \exp\left(-\exp\left(\frac{-x'_{nk}\beta - \sigma_\alpha\alpha}{\sigma_\varepsilon}\right) \cdot z_{nk}^{\frac{1}{\sigma_\varepsilon}}\right).$$

Das auftretende Integral kann nicht analytisch dargestellt werden. Deshalb werden die Log-Likelihood und ihre Ableitungen mit Hilfe der Quadraturformel von Gauß/Hermite approximiert (Bock und Lieberman, 1970, Butler und Moffit, 1982). Angewandt auf das hier vorliegende Problem lautet diese

$$\int_{-\infty}^{\infty} f(\alpha) \cdot \varphi(\alpha) d\alpha = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\alpha) \exp\left(-\frac{\alpha^2}{2}\right) d\alpha \approx \frac{1}{\sqrt{\pi}} \sum_{g=1}^G w_g f(\sqrt{2}h_g),$$

wobei G die Anzahl der verwendeten Stützstellen h_g , und w_g deren Gewichte sind (Stroud und Secrest, 1966). Zur Maximierung der Likelihood wird das Newton-Raphson-Verfahren zusammen mit Techniken zum Auffinden globaler Extrempunkte verwendet (Dennis und Schnabel, 1983). Der ML-Schätzer ist konsistent und asymptotisch normalverteilt unter der Voraussetzung, daß genügend Stützstellen verwendet werden. Da die nötige Anzahl von Stützstellen hier nicht thematisiert werden soll, wird für alle Auswertungen die weitaus ausreichende Zahl von $G = 64$ Stützstellen verwendet. Als Startwerte für diesen iterativen Algorithmus eignen sich i. a. sehr gut die durch die GEE/Buckley-James-Methode (Abschnitt 4.3) ermittelten Schätzergebnisse.

Als Kovariable wird zunächst das in 4.3 als signifikant erkannte Vorhandensein von Fundusvarizen aufgenommen. Zusätzlich erkennt die ML-Methode in jedem Fall die dichotome Variable „Alkoholische Zirrhose“ als signifikant. Zusammen mit der interessierenden Variablen Therapie ergibt sich Tabelle 6. Somit sind Patienten mit Fundusvarizen und nun auch solche mit alkoholischer Zirrhose als Risikogruppen eindeutig erkannt. Der positiv geschätzte Effekt der prophylaktischen Sklerosierung ist auch bei der ML-Methode zu gering, um als signifikant nachgewiesen werden zu können.

Nimmt man zu diesem Modell noch je eine Kovariable der verbleibenden vier (Alter, Geschlecht, Pughpunkte und Bilirubinwert) hinzu, so ergibt sich in keinem dieser Fälle ein zusätzlicher signifikanter Effekt, sondern nur p-Werte über 0,3. Dabei hat die Aufnahme von Alter und Geschlecht wegen der Korrelation mit alkoholischer Diagnose ungenauere Schätzungen zur Folge, während die Berücksichtigung von Pughpunkten und Bilirubinwert die Schätzungen der anderen Parameter kaum beeinflussen.

Ebensowenig ändert sich, wenn man – ähnlich wie bei der Lebensdaueranalyse in Abschnitt 3 – die Tatsache, ob bereits eine Blutung erfolgt ist oder nicht, als (episodenkonstante, aber von Episode zu Episode variierende) Kovariable hinzunimmt. Der Effekt einer bereits erfolgten Blutung wird als intervallverlängernd, aber nicht signifikant geschätzt. Das heißt, wer die erste Blutung erlebt und überlebt, ist im Mittel „gesünder“ als das gesamte Kollektiv zu Beginn der Studie.

Tabelle 6: *Ergebnisse der ML-Schätzung mit Fundusvarizen, Diagnose und Therapie*

		(N)	(L)	(W)
$\hat{\beta}$ [$\hat{\sigma}(\hat{\beta})$]	Konstante	8.65 [0.612]	8.63 [0.612]	9.12 [0.641]
	Fundusvarizen	-2.15 [0.661]	-2.02 [0.652]	-1.86 [0.610]
	Diagnose	-1.27 [0.563]	-1.33 [0.567]	-1.32 [0.570]
	Therapie	0.201 [0.495]	0.221 [0.481]	0.273 [0.460]
$\hat{\sigma}_\alpha$ [$\hat{\sigma}(\hat{\sigma}_\alpha)$]		0.986 [0.450]	1.08 [0.392]	1.05 [0.424]
$\hat{\sigma}_\varepsilon$ [$\hat{\sigma}(\hat{\sigma}_\varepsilon)$]		2.12 [0.289]	1.14 [0.165]	1.37 [0.190]
p-Werte	Fundusvarizen	0.001	0.002	0.003
	Diagnose	0.026	0.020	0.022
	Therapie	0.685	0.647	0.553
l		-353.6	-354.2	-355.2

Eine getrennte Auswertung der beiden Gruppen ohne und mit Fundusvarizen brachte bezüglich keiner der Variablen ein anderes Resultat. Der positive Effekt der prophylaktischen Sklerosierung zeigt sich eher noch in der kleineren Gruppe von Patienten mit Fundusvarizen als in derjenigen ohne Fundusvarizen. Dies liegt wohl in der Tatsache begründet, daß im letzteren Fall bei einer erwarteten Episodenlänge in der Größenordnung von zehn Jahren die Beobachtungszeit in der Größenordnung von zwei Jahren zu kurz ist, um einen zusätzlichen episodenerlängernden Effekt zu erkennen.

Zu Vergleichszwecken mit den Ergebnissen des nächsten Abschnitts sind in Tabelle 7 die Schätzergebnisse des Modells mit den Kovariablen Fundusvarizen, Diagnose, Therapie und Bilirubinwert angegeben. Dabei wird hier jeder Episode derjenige Bilirubinwert zugeordnet, der hinsichtlich des Episodenbeginns in der zeitlich am nächsten liegenden Verlaufsuntersuchung gemessen wurde.

4.5 Die Berücksichtigung der Bilirubinwerte im zeitlichen Verlauf

Die Entwicklung des Bilirubinwerts im zeitlichen Verlauf der Studie ist durch bis zu zehn Verlaufsuntersuchungen pro Patient gut dokumentiert. Die Methoden in 4.3 und 4.4 verwenden diese Informationen nur in relativ wenigen Fällen, nämlich dann, wenn durch eine Blutung eine neue Episode eines Patienten beginnt und in der Nähe des Blutungszeitpunkts ein aktuellerer Wert gemessen wurde. Durch die genaue Berücksichtigung der Entwicklung der Bilirubinwerte erwarten wir einen deutlicheren Effekt und eine genauere Schätzung (vgl. Aydemir, Aydemir und Dirschedl, 1996b).

Tabelle 7: *Ergebnisse der ML-Schätzung mit Fundusvarizen, Diagnose, Therapie und je Episode konstanten Bilirubinwerten*

		(N)	(L)	(W)
$\hat{\beta}$ [$\hat{\sigma}(\hat{\beta})$]	Konstante	8.74 [0.635]	8.71 [0.631]	9.23 [0.657]
	Fundusvarizen	-2.18 [0.664]	-2.05 [0.652]	-1.91 [0.604]
	Diagnose	-1.23 [0.567]	-1.30 [0.569]	-1.27 [0.570]
	Therapie	0.190 [0.495]	0.213 [0.480]	0.258 [0.455]
	Bilirubin	-0.0343 [0.0632]	-0.0354 [0.0590]	-0.483 [0.0529]
$\hat{\sigma}_\alpha$ [$\hat{\sigma}(\hat{\sigma}_\alpha)$]		0.954 [0.455]	1.05 [0.404]	0.979 [0.459]
$\hat{\sigma}_\varepsilon$ [$\hat{\sigma}(\hat{\sigma}_\varepsilon)$]		2.12 [0.289]	1.15 [0.166]	1.38 [0.192]
p-Werte	Fundusvarizen	0.001	0.002	0.002
	Diagnose	0.031	0.024	0.027
	Therapie	0.701	0.658	0.572
	Bilirubin	0.588	0.550	0.364
l		-353.4	-354.0	-354.8

Dazu werden das Modell und das Schätzverfahren aus 4.4 folgendermaßen verallgemeinert (vgl. Petersen, 1986): Die transformierte Verweildauer $y_{nk} = \ln(T_{nk})$ hängt vom zeitvariierenden Kovariablenvektor $x_{nk}(t)$, von β , σ_α , α_n und σ_ε über die Hazardrate

$$\lambda_{nk}(t) = \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} P(t \leq T_{nk} < t + \Delta t \mid T_{nk} \geq t; x_{nk}(t), \alpha_n)$$

ab. Dabei werden die folgenden drei Modellspezifikationen untersucht:

(N)

$$\lambda_{nk}(t) = \frac{1}{\sigma_\varepsilon t} \varphi \left(\frac{\ln t - x'_{nk}(t)\beta - \sigma_\alpha \alpha_n}{\sigma_\varepsilon} \right) \cdot \left(1 - \Phi \left(\frac{\ln t - x'_{nk}(t)\beta - \sigma_\alpha \alpha_n}{\sigma_\varepsilon} \right) \right)^{-1}$$

(Lognormal-Modell),

(L)

$$\lambda_{nk}(t) = \frac{1}{\sigma_\varepsilon} \exp \left(\frac{-x'_{nk}(t)\beta - \sigma_\alpha \alpha_n}{\sigma_\varepsilon} \right) \cdot t^{\frac{1}{\sigma_\varepsilon} - 1} \cdot \left(1 + \exp \left(\frac{-x'_{nk}(t)\beta - \sigma_\alpha \alpha_n}{\sigma_\varepsilon} \right) \cdot t^{\frac{1}{\sigma_\varepsilon}} \right)^{-1}$$

(Loglogistisches Modell) bzw.

(W)

$$\lambda_{nk}(t) = \frac{1}{\sigma_\varepsilon} \exp \left(\frac{-x'_{nk}(t)\beta - \sigma_\alpha \alpha_n}{\sigma_\varepsilon} \right) \cdot t^{\frac{1}{\sigma_\varepsilon} - 1} \quad (\text{Weibull-Modell}).$$

Vom Kovariablenvektor $x_{nk}(t)$ wird angenommen, daß er in (dem beobachteten Teil) der Episode nur endlich viele (J_{nk}) Werte annimmt, also eine Sprungfunktion darstellt mit

$$x_{nk}(t) = x_{nkj} \text{ für } t_{nk,j-1} \leq t < t_{nkj},$$

$j = 1, \dots, J_{nk}$ und den „Sprungzeiten“ $0 = t_{nk0} < t_{nk1} < \dots < t_{nkJ_{nk}} = z_{nk}$. Als Sprungzeiten werden jeweils die Zeitpunkte einer Verlaufsuntersuchung mit aktualisiertem Bilirubinwert, gemessen vom Beginn der Episode an, verstanden.

Die Log-Likelihood hat somit die Gestalt

$$l(\beta, \sigma_\alpha, \sigma_\varepsilon \mid X_1, \dots, X_N, z_{11}, \dots, z_{N, K_N}, \delta_{11}, \dots, \delta_{N, K_N}) = \text{const} + \sum_{n=1}^N \ln \int_{-\infty}^{\infty} \exp \left(\sum_{k=1}^{K_n} \left(\delta_{nk} \cdot \ln \lambda_{nk}(z_{nk}) + \sum_{j=1}^{J_{nk}} \ln \tilde{S}_{nk}(t_{nkj}, x_{nkj}) - \sum_{j=1}^{J_{nk}} \ln \tilde{S}_{nk}(t_{nk,j-1}, x_{nkj}) \right) \right) \cdot \varphi(\alpha) d\alpha$$

mit

(N)

$$\tilde{S}(t, x) = 1 - \Phi \left(\frac{\ln t - x' \beta - \sigma_\alpha \alpha}{\sigma_\varepsilon} \right),$$

(L)

$$\tilde{S}(t, x) = \left(1 + \exp \left(\frac{-x' \beta - \sigma_\alpha \alpha}{\sigma_\varepsilon} \right) \cdot t^{\frac{1}{\sigma_\varepsilon}} \right)^{-1} \quad \text{bzw.}$$

(W)

$$\tilde{S}(t, x) = \exp \left(- \exp \left(\frac{-x' \beta - \sigma_\alpha \alpha}{\sigma_\varepsilon} \right) \cdot t^{\frac{1}{\sigma_\varepsilon}} \right).$$

Das weitere Vorgehen (Approximation der Log-Likelihood und ihrer Ableitungen, Extremwertsuche) entspricht demjenigen in 4.4. Als Startwerte sind in diesem Stadium die Schätzergebnisse aus 4.4 verwendbar.

Der Schätzwert für σ_α nähert sich nun mit zunehmender Anzahl von Iterationen dem Wert 0 an. Man kann dies so interpretieren, daß sich nur die (fälschliche) Annahme konstanter Bilirubinwerte in der Schätzung als unbeobachtete Heterogenität niederschlagen hat, die in das Modell eingefügte Heterogenitätskomponente aber bei dieser Datenkonstellation in Wirklichkeit gar nicht existiert und somit das Modell unnötig verkompliziert. Für diese Argumentation spricht der bereits relativ kleine Wert von $\hat{\sigma}_\alpha$ im Vergleich mit $\hat{\sigma}_\varepsilon$ bei der Annahme konstanter Kovariablen innerhalb einer Episode (Tabellen 6 und 7). Die dort geschätzten Werte ergeben eine Korrelation der logarithmierten Verweildauern eines Patienten von

$$\hat{\rho} = \frac{\widehat{Cov}(y_{nk}, y_{nl})}{\widehat{Var}(y_{nk})} = \frac{\hat{\sigma}_\alpha^2}{\hat{\sigma}_\alpha^2 + \hat{\sigma}_\varepsilon^2} \approx 0.2.$$

Andererseits zeigt ein Blick auf die typische Form der Log-Likelihood im Fall von episodenkonzentrierten Bilirubinwerten (Abbildung 3, durchgezogene Linie),

Tabelle 8: *Ergebnisse der ML-Schätzung mit Fundusvarizen, Diagnose, Therapie und zeitabhängigen Bilirubinwerten*

		(N)	(L)	(W)
$\hat{\beta}$ [$\hat{\sigma}(\hat{\beta})$]	Konstante	8.73 [0.626]	8.70 [0.603]	9.19 [0.629]
	Fundusvarizen	-1.65 [0.484]	-1.56 [0.462]	-1.40 [0.432]
	Diagnose	-1.15 [0.556]	-1.27 [0.557]	-1.26 [0.565]
	Therapie	0.310 [0.452]	0.274 [0.438]	0.379 [0.415]
	Bilirubin	-0.0402 [0.0520]	-0.0389 [0.0435]	-0.534 [0.0341]
$\hat{\sigma}_\alpha$ [$\hat{\sigma}(\hat{\sigma}_\alpha)$]		0 [0.616]	0 [0.803]	0 [0.566]
$\hat{\sigma}_\varepsilon$ [$\hat{\sigma}(\hat{\sigma}_\varepsilon)$]		2.32 [0.279]	1.28 [0.147]	1.49 [0.174]
p-Werte	Fundusvarizen	0.001	0.001	0.001
	Diagnose	0.040	0.024	0.027
	Therapie	0.493	0.534	0.363
	Bilirubin	0.441	0.373	0.120
l		-353.7	-354.4	-355.0

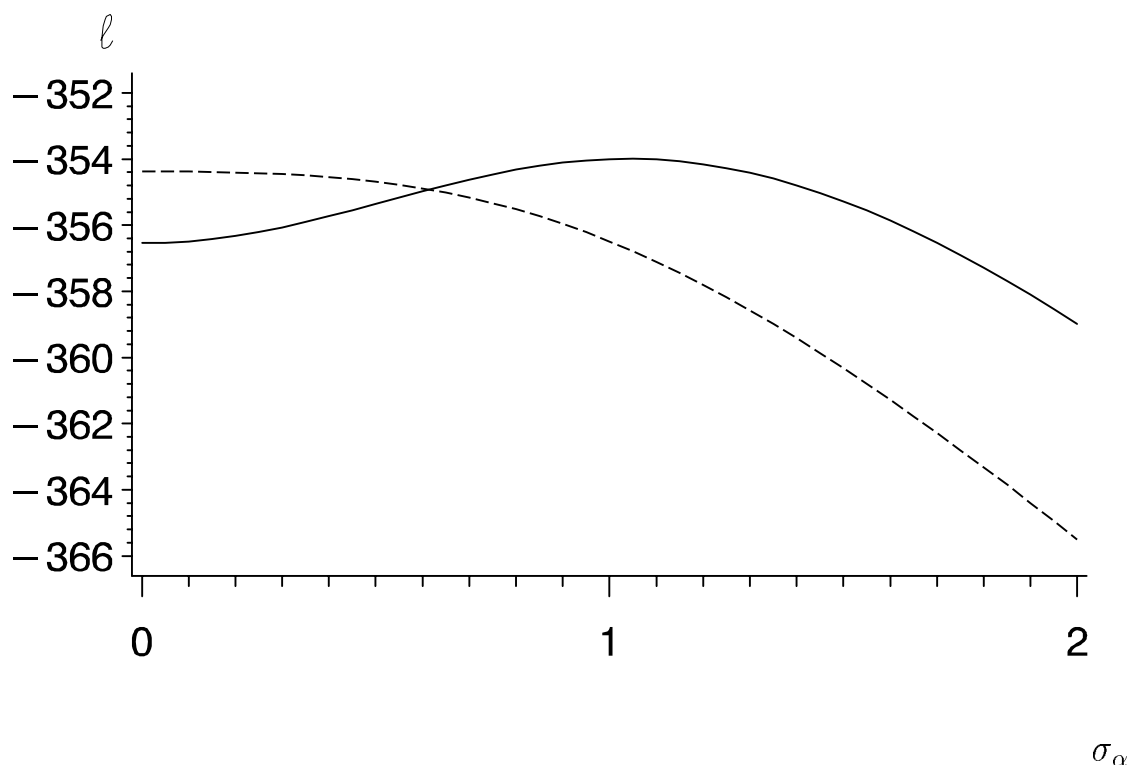
daß ein sehr kleines σ_α , das wegen unzureichender Information (kleiner Stichprobenumfang, hoher Zensierungsgrad) nur ungenau geschätzt werden kann, evtl. nicht von einem tatsächlichen Wert $\sigma_\alpha = 0$ unterschieden werden kann. Zum Vergleich als gestrichelte Linie die Log-Likelihood in Abhängigkeit von σ_α bei zeitvariierenden Bilirubinwerten. Die Log-Likelihood besitzt in $\sigma_\alpha = 0$ in jedem Fall eine horizontale Tangente und nimmt dort in einen Fall ein lokales Minimum, im anderen das globale Maximum an. Der geschätzte Standardfehler von $\hat{\sigma}_\alpha$ zeigt außerdem an, daß die Schätzung auf 0 ungenauer ist als die Schätzungen in 4.4.

Die Schätzung unter Berücksichtigung zeitabhängiger Bilirubinwerte zeigt für Fundusvarizen und alkoholischer Diagnose schwächere Effekte, für Therapie und Bilirubinwert hingegen etwas stärker geschätzte Effekte im Vergleich zu Tabelle 7. Dabei sind die Schätzungen jedoch durchweg genauer geworden.

4.6 Zusammenfassung und Vergleich der Methoden

Zur Untersuchung der Einflußgrößen auf die Zeitabstände von der Aufnahme in die Studie bis zur ersten Blutung bzw. zwischen je zwei Blutungen wurde ein Modell angenommen, bei dem die logarithmierte Verweildauer linear von den bekannten Variablen und einem Frailty-Effekt abhängt. Zur Schätzung des Modells wurden Methoden verwendet, die auf verschiedenen starken Annahmen beruhen und sich durch unterschiedliche numerische Komplexität auszeichnen.

Abbildung 3: *Log-Likelihood des log-logistischen Modells mit Fundusvarizen, Diagnose, Therapie und episodenkonzanten (durchgezogene Linie) bzw. zeitveränderlichen (gestrichelte Linie) Bilirubinwerten in Abhängigkeit von σ_α bei an der Stelle des globalen Maximums festgehaltenen Werten von β und σ_ε*



Während die GEE/Buckley-James-Schätzung keine Annahmen über die Verteilungen der Fehlerterme macht und auch bei Fehlspezifikation der Zusammenhgangsstruktur asymptotisch effiziente Ergebnisse liefert, beruht die ML-Methode entscheidend auf entsprechenden Annahmen. Allerdings hat sich in Simulationsstudien gezeigt, daß auch die ML-Schätzung Robustheitseigenschaften aufweist. Die sehr stabilen Schätzungen im Vergleich zwischen den Spalten (N), (L) und (W) in den Tabellen 6 bis 8 stimmen mit der in Simulationsstudien gewonnenen Erfahrung überein, daß eine Fehlspezifikation der Verteilung der je Episode variierenden Fehlerkomponente keine spürbare Auswirkung zumindest auf die Erwartungstreue der Regressionsparameterschätzung hat.

Ein bedeutender Vorzug der GEE/Buckley-James-Methode liegt bei Zensierungsquoten bis zu einer Größenordnung von 50% in ihrer numerischen Einfachheit. Bei dieser Anwendung mit stärkerer Zensierung ist hingegen die Buckley-James-typische Schleifenbildung entscheidend für einen ebenfalls hohen numerischen Aufwand.

Die Ergebnisse in den Tabellen 5 und 6 können so gedeutet werden, daß im Durchschnitt die Blutungsintervalle eines Patienten mit Fundusvarizen sich zu denen eines Patienten ohne Fundusvarizen etwa wie $e^{-2} : 1 \approx 0.4$ verhalten. Entsprechend ist das

Tabelle 9: *Beobachtete und vorhergesagte Verweildauern von acht Subgruppen*
(F: Fundusvarizen, D: Diagnose, T: Therapie)

F	D	T	Anz. Episoden		⊖ Tage, beob.		Tage, vorhergesagt			
			zens.	unzens.	zens.	unzens.	GEE/BJ	ML(N)	ML(L)	ML(W)
0	0	0	18	4	724	403	4188	5710	5597	4143
0	0	1	21	4	843	352	4317	6981	6981	5444
0	1	0	31	15	594	337	1176	1604	1480	1107
0	1	1	38	18	689	161	1212	1961	1846	1454
1	0	0	1	0	77	-	528	665	742	645
1	0	1	1	1	630	318	545	813	926	847
1	1	0	8	12	128	176	148	187	196	172
1	1	1	2	0	307	-	153	228	245	226
Σ			120	54						

Verhältnis bei alkoholischer Zirrhose versus anderer Diagnose etwa $e^{-1.3} : 1 \approx 0.27$. Der intervallverlängernde Effekt einer prophylaktischen Sklerosierung wird auf einen Faktor von etwa $e^{0.2} : 1 \approx 1.2$ geschätzt. In Tabelle 9 sind für diese durch drei dichotome Variablen entstehenden acht Subgruppen die von den Schätzungen „vorhergesagten“ Verweildauern und die Mittelwerte der Beobachtungen angegeben. Dabei sind die Verweildauern eines Patienten mit K_n Episoden jeweils mit dem Faktor $1/K_n$ gewichtet.

Die Hinzunahme der zeitveränderlichen, aber in diesem Schritt während einer Episode konstant angenommenen Variable „Bilirubinwerte“ ändert zunächst an den Schätzungen der anderen drei Effekte kaum etwas (Tabelle 7). Auch die Schätzungen der Fehlerstreuungen und die Log-Likelihood werden fast nicht beeinflusst. Es wird ein schwacher Zusammenhang von höheren Bilirubinwerten mit kürzeren Blutungsabständen geschätzt, der aber nicht signifikant ist.

Die Berücksichtigung der Veränderungen der Bilirubinwerte während der Episoden hat unter allen drei Modellannahmen zur Folge, daß der negative Zusammenhang mit der Verweildauer etwas deutlicher zum Vorschein kommt, und zudem die Standardfehler geringer geschätzt werden (Tabelle 8). Dies wirkt sich aber auch auf die anderen Effekte aus. Insbesondere wird auch die Wirkung der prophylaktischen Sklerosierung bei niedrigeren Standardfehlern höher geschätzt. Zwar ändert dies bei diesem Datensatz nichts daran, daß der Effekt nicht als signifikant nachgewiesen werden kann, jedoch wird deutlich, daß die verwendeten, verfeinerten Methoden einen Informationsgewinn zur Folge haben können. Das Einbeziehen der Änderungen innerhalb der Episoden ist allerdings wiederum mit hohem numerischen Aufwand verbunden.

5 Diskussion

Am Beispiel einer Studie über Oesophagusvarizen wurden Verfahren zur Überlebenszeitanalyse bei Vorliegen von zeitveränderlichen Kovariablen und Methoden der multivariaten Ereignisanalyse rekurrenter Ereignisse diskutiert. Untersucht wurde die Verweildauer bis zum Auftreten von Varizenblutungen sowie die Überlebenszeit der

Patienten, wobei für letzteres Varizenblutungen einen potentiellen Prognosefaktor darstellen.

Für die multivariate Analyse der Zeit bis zur Varizenblutung wurde ein Accelerated-Failure-Time-Modell mit Frailty-Term verwendet. Die Schätzungen wurden mit einer kombinierten GEE/Buckley-James-Methode und alternativ durch Maximierung der marginalen Likelihood vorgenommen. Mit der GEE/Buckley-James-Methode resultierte lediglich für die Kovariable Fundusvarizen ein auffälliger Einfluß auf die Zeit bis zum Auftreten von Varizenblutungen. Mit der zweiten Schätzmethode, der Maximierung der marginalen Likelihood, zeigte sich neben den Fundusvarizen auch für die Diagnose ein signifikanter Effekt. Sowohl das Vorhandensein von Fundusvarizen als auch die Diagnose einer alkoholischen Zirrhose besaßen demnach einen auffällig verkürzenden Einfluß auf das Intervall bis zur Blutung. Diese Ergebnisse entsprechen auch den Erfahrungen der Mediziner, die allgemein von einem erhöhten Blutungsrisiko bei Patienten mit alkoholischer Zirrhose und Varizen im Magenfundus ausgehen.

In weiterführenden Analysen wurden zur Erklärung der Verweildauer bis zur Blutung zusätzlich in Follow-up Untersuchungen ermittelte Bilirubinwerte einbezogen. Die numerisch aufwendige Berücksichtigung der Bilirubinwerte im zeitlichen Verlauf brachte allerdings keine neuen Erkenntnisse aus der Sicht der Anwendung. Aus dem Blickwinkel der statistischen Modellbildung liegt hier aber ein interessantes Zusammenspiel von unbeobachteter Heterogenität und Berücksichtigung zeitabhängiger Kovariablen vor.

Für die Analyse der Überlebenszeit wurden das lineare Modell von Aalen und semiparametrische Cox-Regressionsmodelle eingesetzt. Varizenblutungen, die in den vorliegenden Daten bei einigen Patienten bis zu dreimal auftraten, wurden als zeitveränderliche Kovariable einbezogen. Da zweite und dritte Varizenblutungen selten vorkamen, wurden unterschiedliche zeitveränderliche Analysen durchgeführt, die entweder nur die erste Blutung, die ersten beiden Blutungen oder alle drei Blutungen berücksichtigten. Der Vergleich der zeitveränderlichen Schätzungen mit fixen Analysen, die die zeitliche Struktur der Blutungen außer acht lassen, zeigte bis auf die Effekthöhe bzw. Einflußstärke der Varizenblutungen kaum Unterschiede. Varizenblutungen hatten in fixen sowie in zeitveränderlichen Schätzungen stets einen signifikant positiven Einfluß auf den Hazard, wobei in fixen Analysen die Effekthöhe geringer ausfiel als in entsprechenden zeitveränderlichen Modellen. Neben Varizenblutungen besaßen höheres Alter und höhere Pugh-Child-Klassifikation einen signifikant ungünstigen Einfluß auf das Überleben. Ferner hatten Frauen gegenüber Männern eine auffällig günstige Prognose, allerdings nur in zeitveränderlichen Analysen mit dem time-dependent Cox-Modell. In Analysen mit dem linearen Modell sowie in fixen Schätzungen mit dem Cox-proportional-hazards-Modell war dagegen kein signifikanter Einfluß des Geschlechts erkennbar.

Interessant war auch der Vergleich der zeitveränderlichen Analysen für unterschiedliche Modellierungen der Varizenblutungen. Wurde nur die erste Blutung der Patienten berücksichtigt, so schien der negative Einfluß auf das Überleben stärker zu sein als bei Einbezug der nachfolgenden Blutungen. Diese Effektschwächung in Modellen, die die zweiten bzw. dritten Blutungen berücksichtigten, wurde dadurch begründet, daß kurz

nach der ersten Blutung bereits viele Patienten verstarben und folglich diejenigen, die die Möglichkeit hatten, ein zweites Mal zu bluten, ein selektiertes Kollektiv gesünderer Patienten bildeten. Vergleichbares konnte auch in den multivariaten Analysen der Verweildauern bis zur Varizenblutung beobachtet werden. Hier zeigte sich bezüglich einer bereits erfolgten Blutung ebenfalls ein intervallverlängernder Trend, d.h. das nach der ersten Blutung noch unter Risiko stehende Kollektiv war gemäß diesen Ergebnissen im Mittel „gesünder“ als das gesamte Studienkollektiv zu Erhebungsbeginn.

Literatur

- Aalen**, Odd O. (1989). A linear regression model for the analysis of life times. *Statistics in Medicine*, 8, 907-925.
- Aalen**, Odd O. (1993). Further results on the non-parametric linear regression model in survival analysis. *Statistics in Medicine*, 12, 1569-1588.
- Altman**, Douglas G. und **De Stavola**, Bianca L. (1994). Practical problems in fitting a proportional hazards model to data with updated measurements of the covariates. *Statistics in Medicine*, 13, 301-341.
- Aydemir**, Sibel, **Aydemir**, Ülker und **Dirschedl**, Peter (1996a). *Das lineare Regressionsmodell von Aalen zur Analyse von Überlebenszeiten unter Berücksichtigung zeitveränderlicher Kovariablen*. Discussion Paper 25, Sonderforschungsbereich 386 der Ludwig-Maximilians-Universität München.
- Aydemir**, Sibel, **Aydemir**, Ülker und **Dirschedl**, Peter (1996b). *Survivalanalysen mit Berücksichtigung der zeitlichen Kovariablenentwicklung in klinischen Studien*. Discussion Paper 44, Sonderforschungsbereich 386 der Ludwig-Maximilians-Universität München.
- Bock**, R. Darrell und **Lieberman**, Marcus (1970). Fitting a response model for n dichotomously scored items. *Psychometrika*, 35 (2), 179-197.
- Buckley**, Jonathan und **James**, Ian (1979). Linear regression with censored data. *Biometrika*, 66, 429-436.
- Butler**, J. S. und **Moffit**, Robert (1982). Notes and comments: A computationally efficient quadrature procedure for the one-factor multinomial probit model. *Econometrica*, 50 (3), 761-764.
- Currie**, Iain D. (1996). A note on Buckley-James estimators for censored data. *Biometrika*, 83, 912-915.
- Dennis**, J. E. jr. und **Schnabel**, R. B. (1983). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Fahrmeir**, Ludwig, **Hamerle**, Alfred und **Tutz**, Gerhard (1996). Regressionsmodelle zur Analyse von Verweildauern. In: Ludwig Fahrmeir, Alfred Hamerle and Gerhard Tutz (Hrsg.): *Multivariate statistische Verfahren*. (2. Aufl.) Berlin: de Gruyter, S. 301-356.

- Hornsteiner, Ulrich und Hamerle, Alfred** (1996). *A combined GEE/Buckley-James method for estimating an Accelerated Failure Time Model of multivariate failure times*. Discussion Paper 47, Sonderforschungsbereich 386 der Ludwig-Maximilians-Universität München.
- Hornsteiner, Ulrich, Hamerle, Alfred und Michels, Paul** (1997). *Parametric versus Nonparametric Treatment of Unobserved Heterogeneity in Multivariate Failure Times*. Discussion Paper 80, Sonderforschungsbereich 386 der Ludwig-Maximilians-Universität München.
- Kaplan, E.L. und Meier, Paul** (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53, 457-481.
- Liang, Kung-Yee und Zeger, Scott L.** (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73, 13-22.
- Mau, Jochen.** (1988). A comparison of counting process models for complicated life histories. *Applied Stochastic Models and Data Analysis*, 4, 283-298.
- Miller, Rupert G. jr.** (1981). *Survival Analysis*. New York: John Wiley & Sons.
- Petersen, Trond** (1986). Estimating fully parametric hazard rate models with time-dependent covariates. *Sociological Methods & Research*, 14, 219-246.
- SAS Institute Inc.** (1989). *SAS/IML Software: Usage and Reference, Version 6*. Cary, NC: SAS Institute.
- SAS Institute Inc.** (1990). *SAS Language: Reference, Version 6*. Cary, NC: SAS Institute.
- Sauerbruch, Tilman, Wotzka, Rudolf, Köpcke, Wolfgang, Härlein, Michael, Heldwein, Walter, Bayerdörffer, Eckehard, Sander, Rainer, Ansari, Hassan, Starz, Irmgard und Paumgartner, Gustav** (1988). Prophylactic sclerotherapy before the first episode of variceal hemorrhage in patients with cirrhosis. *New England Journal of Medicine*, 319, 8-15.
- Spieß, Martin und Hamerle, Alfred** (1995). *Regression models with correlated binary response variables: A comparison of different methods in finite samples*. Discussion Paper 10, Sonderforschungsbereich 386 der Ludwig-Maximilians-Universität München.
- Stroud, A. H. und Secrest, Don** (1966). *Gaussian Quadrature Formulas*. Englewood Cliffs, New Jersey: Prentice-Hall.