



LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

INSTITUT FÜR STATISTIK
SONDERFORSCHUNGSBEREICH 386



Spatz, Hamerle:

Ein Mehr-Zustands-Mehr-Episoden-Modell in diskreter
Zeit zur Analyse klinischer Studien unter
Beruecksichtigung unbeobachteter Heterogenitaet

Sonderforschungsbereich 386, Paper 82 (1997)

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

Projektpartner



Ein Mehr-Zustands-Mehr-Episoden-Modell in diskreter Zeit zur Analyse klinischer Studien unter Berücksichtigung unbeobachteter Heterogenität

Rita Spatz und Alfred Hamerle*

Oktober 1997

Zusammenfassung

Die in der Theorie der Verweildaueranalyse überwiegenden zeitstetigen Ansätze gehen davon aus, daß keine Bindungen existieren. Diese Voraussetzung ist in der Praxis jedoch problematisch, so daß zeitdiskrete Ansätze der Datensituation besser angepaßt sind.

Im medizinischen Kontext steht zumeist die Überlebenszeit von Patienten im Mittelpunkt. Neben diesem Ereignis können jedoch andere sich gegenseitig ausschließende Ereignisse/Zustände (competing risks) in verschiedenen Episoden, wie etwa der Gesundheitszustand im Krankheitsverlauf, von Interesse sein.

In der vorliegenden Arbeit wird ein zeitdiskreter parametrischer Ansatz zur Analyse von competing risks für Mehr-Episoden-Modellen vorgestellt. Angelehnt an die Theorie der generalisierten linearen Modelle wird ein Multinomiales-Logit-Modell zur Modellierung der Hazardrate verwendet. Für die neben den beobachteten Einflußgrößen bestehende unbeobachtete Heterogenität wird eine Normalverteilungsannahme getroffen. Die Maximum-Likelihood-Schätzung wird mittels des Newton-Raphson-Verfahrens durchgeführt, die nötige Integralapproximation erfolgt über die Gauß-Hermite-Quadraturtechnik. Mit der vorgestellten Methode werden Daten von 476 Patienten einer Hirntumorstudie ausgewertet.

*Lehrstuhl für Statistik, Wirtschaftswissenschaftliche Fakultät, Universität Regensburg

1 Einführung

In der Verweildaueranalyse überwiegen Ansätze und Vorschläge für zeitstetige Modelle (Lawless, 1982; Kalbfleisch und Prentice, 1980; Blossfeld, Hamerle und Mayer, 1989; Lancaster, 1990). Allerdings ist es in vielen Fällen nicht möglich, die Zeitpunkte, an denen Ereignisse stattfinden, exakt anzugeben. Auch wenn davon ausgegangen werden kann, daß der Prozeß selbst einen zeitstetigen Verlauf besitzt, ist es dann lediglich möglich, Zeitintervalle (z.B. Monate) anzugeben. Das gleiche trifft auch für das Auftreten von Zensierungen zu. Bei den meisten zeitstetigen Ansätzen tritt infolgedessen das Problem verzerrter Schätzungen (auch "time-aggregation bias" genannt, der um so größer wird, je größer das Zeitintervall ist) auf, das jedoch häufig vernachlässigt wird (Petersen, 1991). Zudem muß davon ausgegangen werden, daß die Anzahl gleicher Beobachtungswerte (ties) der Untersuchungseinheiten sehr hoch wird. Bei der Verwendung von Likelihood-Schätzmethoden würde dies wiederum zu unbrauchbaren Parameterschätzungen führen, zudem bei der Ableitung der asymptotischen Eigenschaften der Parameterschätzungen die Annahme verwendet wird, daß ties mit Wahrscheinlichkeit 0 auftreten. Auch unter dem Gesichtspunkt, daß es bei zeitstetigen Modellen zumeist größere Schwierigkeiten bereitet, zeitvariable Kovariablen einzubinden, ist es unter Umständen vorteilhafter, zeitdiskrete Hazardraten-Modelle anzuwenden.

Wird davon ausgegangen, daß nur ein bestimmtes Ereignis auftreten kann, handelt es sich um ein Ein-Zustands-Modell. In Frage kommende Ereignisse sind z.B. der Einkauf von Schokolade, das Auftreten eines Herzinfarktes oder die Beendigung eines Beschäftigtenverhältnisses. Können jedoch verschiedene Typen von sich gegenseitig ausschließenden Ereignissen eintreten, so wird von einem Competing-Risks- bzw. von einem Mehr-Zustands-Modell gesprochen, etwa wenn verschiedene Todesursachen von Interesse sind (für stetige Competing-Risks-Modelle vgl. Kalbfleisch und Prentice, 1980). Im sozioökonomischen Kontext kann das Ende von Arbeitslosigkeit differenzierter betrachtet werden, indem z.B. unterschieden wird zwischen einem Wechsel in ein Beschäftigtenverhältnis, Rente (dauerhaft aus dem Erwerbsleben ausgeschieden) und Umschulung. Können die Ereignisse (zumindest ein Ereignis) wiederholt auftreten und können damit eventuell mehrere Zustandswechsel stattfinden, so liegt ein Mehr-Episoden-Modell vor. Beim Krankheitsverlauf eines Patienten folgen die Episoden aufeinander. Ebenso sind aber auch parallel ablaufende Verweildauern, z.B. bei verschiedenen Familienmitgliedern, denkbar.

Patienten einer klinischen Studie unterscheiden sich durch verschiedene Kovariablen, die die Hazardraten beeinflussen. Neben den in das gewählte Modell aufgenommenen Kovariablen können auch unbekannte oder nicht erhobene

Einflußgrößen eine Auswirkung auf die Hazardraten besitzen (z.B. psychischer Zustand, Motivation). Um diese individuellen, nicht durch Kovariablen erfaßten Unterschiede zwischen den Patienten, die sogenannte unbeobachtete Heterogenität, in den Modellen berücksichtigen zu können, gibt es mittlerweile theoretisch einige Lösungsansätze. Allerdings beziehen sich diese Lösungsvorschläge fast ausschließlich auf den zeitstetigen Modellansatz. Im zeitdiskreten Fall verwendet Scheike (1997) bei einer parametrischen Modellierung der unbeobachteten Heterogenität die Annahme einer gammaverteilten und Stahl (1996) die einer normalverteilten unbeobachteten Heterogenitätskomponente. Bei beiden wird allerdings nur der Ein-Zustands-Fall behandelt.

Im folgenden werden zunächst im 2. Abschnitt Modelle für competing risks mit mehreren Episoden vorgestellt. Eine Erweiterung des Modells auf ein Modell, bei dem das Vorliegen unbeobachteter Heterogenität berücksichtigt werden kann, erfolgt in Abschnitt 3. Die verwendete Schätzmethode wird im 4. Abschnitt erläutert. Zuletzt werden in Abschnitt 5 die Ergebnisse der Auswertung einer Hirntumorstudie diskutiert.

2 Mehr-Zustands-Mehr-Episoden-Modell

Für die zeitdiskrete Modellierung der Verweildauern einer Episode e wird die Zeit in Intervalle zerlegt: $[a_0 = 0, a_1), \dots, [a_{q_e-1}, a_{q_e}), [a_{q_e}, \infty)$, wobei a_{q_e} das Ende der Episode e angibt. Das Ende des Beobachtungszeitraums ergibt sich aus der Summe über alle a_{q_e} .

Kennzeichne i (mit $i = 1, \dots, N$) die Untersuchungseinheit (auf diesen Index wird zugunsten der Übersichtlichkeit des öfteren verzichtet), j (mit $j = 1, \dots, m$) den Zustand. Die Untersuchungseinheiten können sich hinsichtlich der Anzahl der Episoden ($e = 1, \dots, E_i$) als auch hinsichtlich der Episodenlänge T_e unterscheiden.

Insofern bedeutet $T_e = t_e$, daß im Intervall $[a_{t_e-1}, a_{t_e})$ ein Zustandswechsel stattfindet und damit die e -te Episode beendet ist. Zugleich wird angenommen, daß im gleichen Intervall eine neue Episode beginnt.

Es ergeben sich die zufälligen Übergangszeiten $T_0 = 0 \leq T_1 \leq T_2 \leq \dots$ mit den zugehörigen Zustandsvariablen Y_0, Y_1, Y_2 , usw. (mit $Y_e \in \{1, \dots, m\}$).

Die diskrete Hazardrate, zum t_e -ten Zeitintervall in den j -ten Zustand zu wechseln, ergibt sich zu:

$$\lambda_j(t_e | \mathbf{x}_{t_e}) = P(T_e = t_e, Y_e = j | T_e \geq t_e, \mathbf{x}_{t_e}),$$

wobei mit \mathbf{x}_{t_e} der r -dimensionale Kovariablenvektor zum t_e -ten Zeitintervall gekennzeichnet wird, der die Verweildauer T_e und den Zustand Y_e beeinflusst.

Er kann konstant über alle Zeitintervalle sein oder von Zeitintervall zu Zeitintervall variieren, ist aber innerhalb eines Zeitintervalls konstant. Insofern kann dieser Informationen über die Vergangenheit des Prozesses (z.B. Informationen über die Anzahl bisher erlebter Episoden etc.) enthalten.

Die Gesamthazardrate ergibt sich damit zu

$$\lambda(t_e|\mathbf{x}_{t_e}) = \sum_{j=1}^m \lambda_j(t_e|\mathbf{x}_{t_e})$$

und stellt die bedingte Wahrscheinlichkeit dar, den Zustand y_{e0} , den Startzustand der e -ten Episode, im t_e -ten Zeitintervall zu verlassen (gegeben dieses Intervall erreicht zu haben).

Die Wahrscheinlichkeit, nicht zu wechseln, wird mit $1 - \lambda(t_e|\mathbf{x}_{t_e})$ angegeben, so daß sich die diskrete Survivorfunktion berechnet zu

$$S(t_e|\mathbf{x}_{t_e}) = P(T_e > t_e|\mathbf{x}_{t_e}) = \prod_{s=0}^{t_e} (1 - \lambda(s|\mathbf{x}_s)).$$

Im Rahmen der generalisierten linearen Modelle wird die Hazardrate allgemein über die lokale Linkfunktion h_j für den Zustand j im Zeitintervall $[a_{s-1}, a_s)$ (mit $s = 0, \dots, t_e$) und dem linearen Prediktor $\eta_{sj} = \mathbf{z}'_{sj}\beta$ definiert:

$$\lambda_j(s|\mathbf{x}_s) = h_j(\eta_{sj}) = h_j(\mathbf{z}_s\beta),$$

wobei \mathbf{z}_s eine von den Kovariablen \mathbf{x}_s und der Zeit s abhängige Designmatrix ist (mit den Zeilen \mathbf{z}_{sj}) und β der zu schätzende Parametervektor.

Zur Modellierung der zustandsspezifischen Hazardraten bieten sich das Multinomiale-Logit-Modell und das gruppierte Cox-Modell an. Denkbare wären aber auch (gestützt auf die Theorie der generalisierten linearen Modelle) sequentielle und kumulative Modelle (vgl. Fahrmeir und Tutz, 1994, S. 332). Wie Stahl (1996) für den Ein-Zustands-Fall zeigt, unterscheiden sich das Logit-Modell und das gruppierte Cox-Modell v.a. für kurze Intervalle nicht wesentlich (vgl. auch Thompson, 1977). Im weiteren wird explizit auf das Multinomiale-Logit-Modell (vgl. Allison, 1982) eingegangen:

$$\begin{aligned} \lambda_{sj} =: \lambda_j(s|\mathbf{x}_s) &= \frac{\exp(\eta_{sj})}{1 + \sum_{k=1}^m \exp(\eta_{sk})} \\ &= \frac{\exp(\beta_{0sj} + \mathbf{x}'_s\beta_{sj})}{1 + \sum_{k=1}^m \exp(\beta_{0sk} + \mathbf{x}'_s\beta_{sk})}, \end{aligned}$$

wobei β_{sj} (mit $j = 1, \dots, m$) die zustandsspezifische Gewichtung der Kovariablen darstellt. Dieser Ansatz führt v.a. bei vielen Intervallen zu einer immensen Anzahl von zu schätzenden Parametern und infolgedessen eventuell dazu, daß

der Maximum-Likelihood-Schätzer nicht mehr berechnet werden kann. Eine parametersparsamere Alternative ist die Modellierung der Baselineparameter $\beta_{0s1}, \dots, \beta_{0sm}$ mit $s = 0, \dots, \sum_{e=1}^E t_e$ über eine Polynomapproximation (z.B. der Form $\beta_{0sj} = \beta_{01j} + s\beta_{02j} + s^2\beta_{03j}$) und eine zeitkonstante Modellierung des Parametervektors $\beta_{js} := \beta_j$. Die entsprechende Designmatrix hat folgende Form:

$$\mathbf{z}_s = \begin{pmatrix} 1 & s & s^2 & \mathbf{x}' & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 1 & s & s^2 & \mathbf{x}' & 0 & \dots & 0 \\ \vdots & & & & & & & & & & \vdots \\ 0 & & & \dots & & & & & 0 & 1 & s & s^2 & \mathbf{x}' & 0 \end{pmatrix}$$

Weitere Vorschläge zur Modellierung zeitveränderlicher Effekte sind bei Yamaguchi (1993) und Hastie und Tibshirani (1993) zu finden. Möglich wäre unter anderem, die zustandsspezifischen Baselineparameter über mehrere Zeitintervalle konstant zu halten.

Ganz allgemein wird im weiteren der zu schätzende Parametervektor β mit $\beta = (\beta_1, \dots, \beta_p)'$ bezeichnet, wobei p entsprechend der gewählten Modellierung von r , m und q abhängt. Werden z.B. die Parameter zeitkonstant aber zustandsspezifisch modelliert, so ergibt sich $p = (r + 1)m$.

Typisch bei der Analyse von Verweildauern ist das Auftreten von Zensierungen. Wird etwa das Ende des Beobachtungszeitraumes erreicht, während die Verweildauer der Untersuchungseinheit noch nicht beendet ist, oder steht die Untersuchungseinheit (UE) aus anderen Gründen (z.B. Umzug) nicht mehr zur Verfügung und die Verweildauer kann nicht mehr ermittelt werden, so spricht man von rechts zensierten Daten. Ausgegangen wird von zufälligen Zensierungen, die zu Beginn eines Intervalls erfolgen und die unabhängig von den Verweildauern sind. Erfolgt eine Zensierung, so wird dies durch eine individuen- und episodenspezifische Indikatorvariable gekennzeichnet:

$$\delta_{ie} = \begin{cases} 0 & \text{falls in der } e\text{-ten Episode UE } i \text{ zensiert wird} \\ 1 & \text{sonst} \end{cases}$$

Werden aufeinanderfolgende Verweildauern betrachtet, so wird üblicherweise nur die letzte Episode zensiert sein. Werden hingegen parallel ablaufende Prozesse betrachtet, so kann prinzipiell jede Episode zensiert sein.

3 Mehr-Zustands-Mehr-Episoden-Modell mit Berücksichtigung unbeobachteter Heterogenität

Da in der angewandten Forschung selten alle relevanten Kovariablen erhoben und in dem jeweiligen Modell aufgenommen werden, muß fast immer vom Vorliegen einer unbeobachteten Heterogenität ausgegangen werden. Wird sie nicht berücksichtigt, so führt dies in den meisten Fällen zu einer abnehmenden Hazardrate, tendenziell in Richtung negativer Zeitabhängigkeit (vgl. hierzu Wangler, 1997).

Bei einer parametrischen Modellierung unbeobachteter Heterogenität wird das bisherige Modell um eine Zufallsvariable erweitert, so daß von den generalisierten linearen Modellen (GLM) auf die generalisierten linearen Modelle mit zufälligen Effekten (GLRCM) übergegangen wird (vgl. Fahrmeir und Tutz, 1994, Kapitel 7):

$$\lambda_j(t_e | \mathbf{x}_{t_e}, b) = P(T_e = t_e, Y_e = j | T_e \geq t_e, \mathbf{x}_{t_e}, b).$$

Von der unabhängig und identisch normalverteilten Heterogenitätskomponente b (mit $E(b) = 0$ und $Var(b) = \sigma^2$) wird angenommen, daß sie zwar individualspezifisch, jedoch für jeden Zustand, für jedes Zeitintervall und damit auch für jede Episode gleich ist. Zudem wird von einer Unabhängigkeit der beobachteten Kovariablen und der unbeobachteten Heterogenitätskomponente ausgegangen.

Über die Reparametrisierung $b = \sigma a$ wird das Multinomiale Logitmodell im linearen Prediktor zu $\eta_{t_e j} = \beta_{0t_e j} + \mathbf{x}'_{t_e} \beta_j + \theta a$ erweitert (für die Standardabweichung σ wird der Parameter θ geschätzt), so daß

$$\begin{aligned} \lambda_j(t_e | \mathbf{x}_{t_e}, a) &= \frac{\exp(\eta_{t_e j})}{1 + \sum_{k=1}^m \exp(\eta_{t_e k})} \\ &= \frac{\exp(\beta_{0t_e j} + \mathbf{x}'_{t_e} \beta_j + \theta a)}{1 + \sum_{k=1}^m \exp(\beta_{0t_e k} + \mathbf{x}'_{t_e} \beta_k + \theta a)}. \end{aligned}$$

4 Maximum-Likelihood-Schätzung

Die unbekanntesten festen Parameter β und θ werden über die Maximierung der marginalen Log-Likelihood mittels iterativer Verfahren geschätzt. Für den Fall, daß die Schätzung ohne Berücksichtigung unbeobachteter Heterogenität

erfolgt, ist der Likelihoodbeitrag der i -ten Untersuchungseinheit gegeben mit

$$L_i(\beta) = \prod_{e=1}^{E_i} \prod_{s=1}^{t_e-1+\delta_{ie}} \left[\left(\prod_{\substack{j=1 \\ j \neq y_{ie0}}}^m \lambda_{sj}(z_{iesj})^{y_{iesj}} \right) \left(1 - \sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m \lambda_{sj}(z_{iesj}) \right)^{y_{ies,m+1}} \right]$$

mit

$$y_{iesj} = \begin{cases} 1 & \text{falls in der } e\text{-ten Episode UE } i \text{ im Intervall } [a_{s-1}, a_s) \\ & \text{in Zustand } j \text{ wechselt} \\ 0 & \text{sonst} \end{cases}$$

und $y_{ies,m+1} = 1 - \sum_{j=1}^m y_{iesj}$. Der Startzustand wird mit y_{ie0} gekennzeichnet, z_{iesj} ist die j -te Zeile der Designmatrix \mathbf{z}_s für das s -te Zeitintervall in der e -ten Episode der i -ten Untersuchungseinheit.

Wird hingegen das Vorliegen unbeobachteter Heterogenität berücksichtigt, so ist

$$L_i(\beta, \theta) = \int \prod_{e=1}^{E_i} \prod_{s=1}^{t_e-1+\delta_{ie}} \left[\left(\prod_{\substack{j=1 \\ j \neq y_{ie0}}}^m \lambda_{js}(\mathbf{x}_s, a)^{y_{iesj}} \right) \cdot \left(1 - \sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m \lambda_{js}(\mathbf{x}_s, a) \right)^{y_{ies,m+1}} \right] g(a) da,$$

wobei $g(a)$ die Dichte von a bezeichnet.

Diese Likelihood enthält ein im allgemeinen analytisch nicht lösbares Integral. Um sie zumindest approximativ zu erhalten, wird, da die Dichte g als normalverteilt angenommen wird, die Gauss-Hermite-Quadraturtechnik verwendet (vgl. Davis und Rabinowitz, 1975 und 1984, Fahrmeir und Tutz, 1994). Die dazu benötigten Stützstellen werden mit d_l ($l = 1, \dots, g$) und die Gewichte der Stützstellen mit w_l bezeichnet. Zudem wird $\kappa_l := d_l \theta \sqrt{2}$ gesetzt.

Die mit dieser Integralapproximation erhaltene Likelihood lautet

$$L_i(\beta, \theta) = \sum_{l=1}^g \frac{w_l}{\sqrt{\pi}} \prod_{e=1}^{E_i} \prod_{s=1}^{t_e-1+\delta_{ie}} \left[\left(\prod_{\substack{j=1 \\ j \neq y_{ie0}}}^m \lambda_j(s|z_{iesj}, d_l)^{y_{iesj}} \right) \cdot \left(1 - \sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m \lambda_j(s|z_{iesj}, d_l) \right)^{y_{ies,m+1}} \right]$$

Beim Multinomialen-Logit-Modell ist dann

$$L(\beta, \theta) = \prod_{i=1}^N \sum_{l=1}^g \frac{w_l}{\sqrt{\pi}} \prod_{e=1}^{E_i} t_e^{-1+\delta_{ie}} \prod_{s=1}^m \left\{ \prod_{\substack{j=1 \\ j \neq y_{ie0}}}^m \left(\frac{\exp(z'_{iesj}\beta + \kappa_l)}{1 + \sum_{\substack{k=1 \\ k \neq y_{ie0}}}^m \exp(z'_{iesk}\beta + \kappa_k)} \right)^{y_{iesj}} \right. \\ \left. \cdot \left(1 - \sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m \frac{\exp(z'_{iesj}\beta + \kappa_l)}{1 + \sum_{\substack{k=1 \\ k \neq y_{ie0}}}^m \exp(z'_{iesk}\beta + \kappa_l)} \right)^{y_{ies,m+1}} \right\}$$

Zur Schätzung der Parameter wird das Newton-Raphson-Verfahren verwendet. Zur Verbesserung des Konvergenzverhaltens wurden (entsprechend Dennis und Schnabel, 1983) Änderungen im Schätzalgorithmus vorgenommen, falls das Vorliegen unbeobachteter Heterogenität berücksichtigt wird. Der Scorevektor $S(\beta)$ und die Hessematrix $H(\beta)$, die für die Schätzgleichung

$$\hat{\beta}^{(k+1)} = \hat{\beta}^{(k)} - H(\hat{\beta}^{(k)})^{-1} S(\hat{\beta}^{(k)})$$

benötigt werden, sind mit und ohne Berücksichtigung unbeobachteter Heterogenität im Anhang explizit dargestellt.

5 Anwendung: Therapiestudie über hochmaligne Hirntumore

5.1 Studienbeschreibung

Die in dieser Arbeit analysierte multizentrische Therapiestudie über hochmaligne Hirntumore (auch als Gliomstudie bekannt) wurde im Zeitraum von 1983 bis 1989 durchgeführt. Dabei wurden maligne supratentorielle Gliome bei Erwachsenen postoperativ mit einer Strahlen- und einer Chemotherapie (BCNU und BCNU+VM26) behandelt.

Es wurden insgesamt 501 Patienten nach der operativen Entfernung der malignen Gliome in zwei randomisierte Therapiegruppen eingeteilt und erhielten je nach Gruppenzugehörigkeit verschiedene Chemotherapien. Die eine Gruppe wurde mit BCNU, die andere mit BCNU+VM26 behandelt; alle Patienten erhielten unabhängig von der Gruppenzuteilung eine einheitliche Strahlentherapie. Von 25 Patienten fehlten bei Studieneintritt verschiedene Größen. Im

Vergleich zur Gesamtanzahl ist die Zahl der Patienten mit missing values gering, so daß sich die durchgeführte Analyse auf 476 vollständige Datensätze (Monatsdaten) stützt.

- Zielzustände

Bei den bisher durchgeführten Analysen (Ulm et al., 1989, Aydemir et al., 1996) stand lediglich der Einfluß verschiedener Faktoren auf die Überlebenszeit im Mittelpunkt des Interesses. Allerdings wird die Fragestellung mit dem alleinigen Kriterium "Tod" den Patienten nicht gerecht, da für diese nicht nur das Leben, sondern auch die Lebensqualität von entscheidender Bedeutung ist. Bei dem hier weiter verfolgten Ansatz wird neben dem Ereignis "Tod" der Befindlichkeitszustand, der sogenannte Karnofskyindex, betrachtet. Dieser wird in drei Kategorien eingeteilt:

- 81-100% sehr guter körperlicher Zustand
- 71-80% guter körperlicher Zustand
- <71% schlechter körperlicher Zustand

Es ergeben sich infolgedessen vier miteinander konkurrierende Zustände (mit "Zustand Tod" als absorbierendem Zustand) für ein Competing-Risks-Modell. Die (maximal 5) Episoden werden durch den Wechsel eines Patienten zwischen den verschiedenen Befindlichkeitszuständen charakterisiert.

- Einflußgrößen

Wie sich bei der Analyse von Aydemir et al. (1996) zeigte, ist gerade die zeitliche Entwicklung der Einflußgrößen bei dieser Studie von besonderer Bedeutung. Allerdings bereitet die Berücksichtigung zeitvariierender Kovariablen bei Verwendung zeitstetiger Modelle größere Probleme. Aydemir et al. (1996) konnten lediglich eine Kovariable mit ihren sich im Verlauf der Studie ändern den Ausprägungen mitführen; alle anderen Einflußgrößen gingen mit den beim Eintritt in die Studie beobachtbaren Werten ein. Bei dem hier verwendeten zeitdiskreten Modellansatz ist die Beachtung von Veränderungen der Kovariablen im Beobachtungszeitraum kein größeres Problem.

Folgende Kovariablen gehen zeitkonstant in die Schätzung ein:

- Therapiegruppe (0=BCNU, 1=BCNU+VM26)
- Alter in Effektkodierung (Alter1=16-50 Jahre, Alter2=51-60 Jahre, Referenzkategorie: >60 Jahre)
- Biopsie (0=nein, 1=ja)

- Totalresektion (0=nein, 1=ja)
- Malignität in Effektkodierung (Malign1=Grad 3, Malign2=Grad 4 - bösartiger als Grad 3, Referenzkategorie: Grad 3-4)
- Bewußtseinsstörung (0=nein, 1=ja)

Folgende Kovariablen können zeitvariabel berücksichtigt werden (vgl. hinsichtlich deren Struktur Aydemir et al., 1996)

- Cortison (0=nein, 1=ja)
- epileptische Anfälle (0=nein, 1=ja)
- organisches Psychosyndrom (0=nein, 1=ja)

• Vorgehen und verwendete Modelle

Im ersten Schritt werden alle Kovariablen zeitkonstant, also mit ihren Ausprägungen bei Studieneintritt, in der Analyse verwendet. Dann werden die Kovariablen, bei denen zeitliche Verläufe zur Verfügung stehen, mit ihren zeitintervallrelevanten Ausprägungen eingebunden.

Es wird ein Mehr-Zustands-Mehr-Episoden-Modell verwendet, bei dem für die zustandsspezifischen Hazardraten ein Multinomiales-Logit-Modell verwendet wird. Die zu schätzenden Parameter werden zeitpunktunabhängig spezifiziert:

$$\lambda_j(s|x_s) = \frac{\exp(\beta_{0j} + x'_s \beta_j)}{1 + \sum_{k=1}^m \exp(\beta_{0k} + x'_s \beta_k)},$$

mit dem r -dimensionalen Kovariablenvektor x_s im Zeitintervall s ($s = 1, \dots, t_e$) und dem zugehörigen Parametervektor $\beta'_j = (\beta_{1j}, \dots, \beta_{rj})$.

Um zum einen die Anzahl der zu schätzenden Parameter nicht noch weiter zu erhöhen (mit der bisherigen Modellierung sind bereits 48 Parameter zu schätzen) und zum anderen die Zeitdauer s zu berücksichtigen, wird anstelle des linearen zustandsspezifischen Prediktors $\beta_{0j} + x'_s \beta_j$ die Linearkombination

$$\beta_{01j} + s \cdot c \cdot \beta_{02j} + x'_s \beta_j$$

mit $c = 0.1$ verwendet, da ansonsten größere numerische Probleme auftreten. Mit der Verweildauer s wird zunächst die Zeit, die ein Patient bereits in der e -ten Episode verbringt, in das Modell eingebunden, dann die Zeit, die seit seinem Eintritt in der Studie vergangen ist. Zuletzt werden, um die Vergangenheit des Prozesses mitzuberechnen, exemplarisch die Kovariablen um den Zustand in der vorangegangenen Episode ergänzt. Die Daten werden sowohl ohne als auch mit Berücksichtigung unbeobachteter Heterogenität ausgewertet. Die Ergebnisse der Analysen sind im Anhang angeführt.

5.2 Ergebnisse

Aus theoretischer Sicht ist es von Interesse, inwieweit sich die Berücksichtigung unbeobachteter Heterogenität und die Verwendung der zeitlichen Verläufe von Kovariablen bei den Auswertungen bemerkbar machen. Zudem soll die Beantwortung der erweiterten Fragestellung Aufschluß darüber geben, wie sich die verschiedenen Größen nicht wie bisher nur auf den Tod, sondern auch auf die verschiedenen Befindlichkeitszustände auswirken.

- Mit und ohne Berücksichtigung unbeobachteter Heterogenität

Die Ergebnisse der Analysen mit Berücksichtigung unbeobachteter Heterogenität enttäuschen: die Standardabweichung der unbeobachteten Heterogenität ist zu klein, um merkliche Unterschiede der Auswertungsergebnisse bei einer Analyse mit oder ohne Berücksichtigung unbeobachteter Heterogenität zu verursachen.

Ohne unbeobachtete Heterogenität werden die Varianzen der Schätzungen leicht unterschätzt. Lediglich für den Fall, daß alle Kovariablen zeitkonstant (mit ihren Ausprägungen bei Studieneintritt) verwendet werden, führt dies zusammen mit einem betragsmäßig größeren Schätzwert dazu, daß für eine Totalresektion ein Einfluß auf einen guten körperlichen Zustand (71 – 80%) ermittelt wird (p-value: 0.0424, vgl. Tabelle 1 im Anhang).

Für das einfache Modell mit zeitvariablen Kovariablen wird die Standardabweichung der unbeobachteten Heterogenität bereits geringer und die zugehörige Varianz größer geschätzt (vgl. Tabelle 4). Bei allen übrigen Modellen, die zeitvariable Kovariablen und die Verweildauer bei der Baselinemodellierung berücksichtigen, wird die Standardabweichung fast gleich Null geschätzt, so daß ein Vergleich sich erübrigt.

Wie die Ergebnisse der Modelle mit zeitvariablen Einflußgrößen zeigen, ist die bei den Modellen mit zeitkonstanten Kovariablen bemerkbare unbeobachtete Heterogenität auf die nicht genutzte Information der zeitlichen Struktur entsprechender Größen zurückzuführen.

- Zeitkonstante versus zeitvariable Verwendung der Kovariablen

Wenn die Kovariable Epileptische Anfälle zeitkonstant eingebunden wird, so wird ihr negativer Einfluß auf das Überleben nicht erkannt. Wie ein Vergleich von Tabelle 2 und 4 zeigt (Auswertung mit Berücksichtigung unbeobachteter Heterogenität) wird die Varianz mit 0.0498 zeitkonstant größer geschätzt als die Varianz der zeitvariablen Größe mit 0.0180. Die Parameterschätzungen ändern zudem ihr Vorzeichen (die nicht signifikanten Parameterschätzungen

besitzen im zeitkonstanten Fall ein negatives Vorzeichen). Zudem werden epileptischen Anfällen im zeitkonstanten Modell fälschlicherweise ein negativer Einfluß auf einen guten körperlichen Zustand (71 – 80%) zugestanden, da ein größerer Schätzwert (zeitkonstant: -0.4939, zeitvariabel: 0.009) ermittelt wird. Die Einnahme von Cortison wirkt sich in jedem Fall lebensverkürzend aus. Werden jedoch Änderungen bezüglich der Einnahme (ja/nein) von Cortison im Verlauf der Studie bei den Schätzungen ignoriert, so wird dessen Einfluß etwas unterschätzt (zeitkonstant: 0.2736, zeitvariabel: 0.8364). Eine positive Beeinflussung auf einen guten körperlichen Zustand (71 – 80%) läßt sich nur dann nachweisen, wenn bei der Modellierung der Baselineparameter die Verweildauer (in Form von Polynomen) nicht berücksichtigt wird. Für alle anderen Zustände lassen sich keine signifikanten Effekte zeigen.

Falls die zeitliche Struktur der Kovariable organisches Psychosyndrom nicht berücksichtigt wird, kommt es zu einer gravierenden Fehleinschätzung hinsichtlich ihres Einflusses. Bei etwa gleichen Varianzen wird aufgrund unterschätzter Effekte zum einen ein positiver Einfluß auf einen schlechten körperlichen Zustand (zeitkonstant: 0.1333, zeitvariabel: 0.5578) und zum anderen auch der lebensverkürzende Einfluß (zeitkonstant: -0.0158, zeitvariabel: 0.5199) und der negative Einfluß auf einen sehr guten körperlichen Zustand (zeitkonstant: -0.0645, zeitvariabel: -0.7523) nicht erkannt.

- Zustand der vorangegangenen Episode als weitere Einflußgröße

Wird die Vergangenheit des Prozesses in Form eines erweiterten Kovariablenvektors, der zusätzlich den Zustand der vorangegangenen Episode enthält, mit einbezogen (vgl. Tabelle 7), so werden bei etwa gleichbleibenden Varianzen die Schätzungen der meisten Parameter kleiner. Auffallend ist, daß sich lediglich die signifikanten Schätzwerte für die Malignität leicht erhöhen.

Der sich ergebende Effekt der Kovariablen “Zustand der vorangegangenen Episode“ ist trivial: sich in einem schlechten körperlichen Zustand zu befinden wirkt sich negativ auf das Überleben (positive Schätzung für den Zustand Tod) und auf das Erreichen eines sehr guten körperlichen Zustandes (81 – 100%) aus.

- Interpretation der in allen Modellen zeitkonstanten Kovariablen
Übereinstimmend mit Aydemir et al. (1996) hat die Therapie keinen signifikanten Einfluß auf das Überleben und auf einen Wechsel in die verschiedenen Befindlichkeitszustände. Das gleiche gilt für das Vorliegen von Bewußtseinsstörungen und falls bei Eintritt in die Studie der Patient zwischen 51 und 60 Jahre (Alter²) ist.

Wird eine Biopsie vorgenommen, so wird für den Zustand “Tod“ ein posi-

tiver Zusammenhang nachgewiesen (Schätzwerte liegen zwischen 0.7592 und 0.8395).

Ein Tumor vom Grad 3 (Malign1) wirkt sich auf die verschiedenen körperlichen Zustände und den Tod negativ aus, falls alle übrigen Kovariablen zeitkonstant verwendet werden (Tabellen 1 und 2). Der negative Einfluß auf einen sehr guten körperlichen Zustand läßt sich allerdings nicht mehr nachweisen, wenn die Kovariablen epileptische Anfälle, Cortison und organisches Psychosyndrom zeitvariabel eingebunden werden (Tabellen 3-7). Gehen zusätzlich die Verweildauern bei der Modellierung der Baselineparameter ein, so läßt sich nur noch der signifikant positive Einfluß auf das Überleben (negativer Schätzwert für den Zustand Tod) nachweisen. Ein bösartigerer Tumor (Malign2) besitzt entsprechend einen lebensverkürzenden Effekt, für die verschiedenen Befindlichkeitszustände lassen sich jedoch keine signifikanten Schätzwerte ermitteln. Das Alter zwischen 16 und 50 Jahre (Alter1) wirkt sich negativ auf einen schlechten körperlichen Zustand und den Tod aus. Für die anderen Zustände läßt sich kein signifikanter Einfluß nachweisen. Lediglich wenn die Verweildauer bei der Modellierung der Baselineparameter zum Tragen kommt, zeigt sich ein positiver Effekt auf einen sehr guten körperlichen Zustand. Der Effekt der Kovariable Alter1 sollte jedoch unter Vorbehalt interpretiert werden, da diese vom Charakter her zeitvariable Einflußgröße zeitkonstant verwendet wird.

Wird eine Totalresektion durchgeführt, so läßt sich ein Einfluß auf den Tod und auf einen Wechsel in den mittleren Gesundheitszustand (71 – 80%) nicht nachweisen. Ein sehr guter körperlicher Zustand wird hingegen durch eine Totalresektion positiv beeinflusst. Ein negativer Effekt kann für den Zustand < 71% gezeigt werden. Dieser Effekt läßt sich allerdings nicht mehr nachweisen, wenn die Verweildauer in der Studie oder wenn die vorgangene Episode berücksichtigt wird.

Anhang

A Formeln

A.1 Ohne Berücksichtigung unbeobachteter Heterogenität

$$\begin{aligned}
\ln L(\beta) &= \sum_{i=1}^N \sum_{e=1}^{E_i} \sum_{s=1}^{t_e-1+\delta_{ie}} \left[\sum_{j=1}^m y_{iesj} \ln(\lambda_j(s|z_{iesj})) + y_{ies,m+1} \ln\left(1 - \sum_{j=1}^m \lambda_j(s|z_{iesj})\right) \right] \\
&= \sum_{i=1}^N \sum_{e=1}^{E_i} \sum_{s=1}^{t_e-1+\delta_{ie}} \left[\sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m y_{iesj} \ln \left(\frac{\exp(z'_{iesj}\beta)}{1 + \sum_{\substack{k=1 \\ k \neq y_{ie0}}}^m \exp(z'_{iesk}\beta)} \right) \right. \\
&\quad \left. + y_{ies,m+1} \ln \left(1 - \sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m \frac{\exp(z'_{iesj}\beta)}{1 + \sum_{\substack{k=1 \\ k \neq y_{ie0}}}^m \exp(z'_{iesk}\beta)} \right) \right] \\
S(\beta) = \begin{pmatrix} \frac{\partial \ln L}{\partial \beta_1} \\ \vdots \\ \frac{\partial \ln L}{\partial \beta_p} \end{pmatrix} &= \begin{pmatrix} \sum_{i=1}^N \sum_{e=1}^{E_i} \sum_{s=1}^{t_e-1+\delta_{ie}} \sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m z_{iesj}^1 \left(y_{iesj} - \frac{\exp(z'_{iesj}\beta)}{1 + \sum_{\substack{k=1 \\ k \neq y_{ie0}}}^m \exp(z'_{iesk}\beta)} \right) \\ \vdots \\ \sum_{i=1}^N \sum_{e=1}^{E_i} \sum_{s=1}^{t_e-1+\delta_{ie}} \sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m z_{iesj}^p \left(y_{iesj} - \frac{\exp(z'_{iesj}\beta)}{1 + \sum_{\substack{k=1 \\ k \neq y_{ie0}}}^m \exp(z'_{iesk}\beta)} \right) \end{pmatrix}
\end{aligned}$$

mit z_{iesj}^a : a -tes Element der j -ten Zeile der Designmatrix z_i der i -ten UE für das s -te Zeitintervall in der e -ten Episode.

$$H(\beta) = \frac{\partial^2 l}{\partial \beta \partial \beta'} = \begin{pmatrix} \frac{\partial^2 l}{\partial \beta_1^2} & \frac{\partial^2 l}{\partial \beta_1 \partial \beta_2} & \cdots & \frac{\partial^2 l}{\partial \beta_1 \partial \beta_p} \\ \frac{\partial^2 l}{\partial \beta_2 \partial \beta_1} & & & \\ \vdots & & & \\ \frac{\partial^2 l}{\partial \beta_p \partial \beta_1} & \cdots & & \frac{\partial^2 l}{\partial \beta_p^2} \end{pmatrix}$$

Mit $(a, b = 1, \dots, p)$ z_{iesj}^a - a -tes Element der j -ten Zeile der Designmatrix z_{is} der e -ten Episode)

$$\frac{\partial^2 l}{\partial \beta_a \partial \beta_b} = \sum_i^N \sum_{e=1}^{E_i} \sum_{s=1}^{t_e-1+\delta_{ie}} \sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m \frac{-z_{iesj}^a z_{iesj}^b \exp(z'_{iesj} \beta)}{\left(1 + \sum_{\substack{k=1 \\ k \neq y_{ie0}}}^m \exp(z'_{iesk} \beta)\right)^2}$$

A.2 Mit Beachtung einer unbeobachteten Heterogenität

$$\begin{aligned} \ln L(\beta', \theta) &= \sum_{i=1}^N \ln \left\{ \sum_{l=1}^g \frac{w_l}{\sqrt{\pi}} \exp \left[\sum_{e=1}^{E_i} \sum_{s=1}^{t_e-1+\delta_{ie}} \left(\sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m y_{iesj} \ln \left(\frac{\exp(z'_{iesj} \beta + \kappa_l)}{1 + \sum_{\substack{k=1 \\ k \neq y_{ie0}}}^m \exp(z'_{iesk} \beta + \kappa_l)} \right) \right. \right. \right. \\ &\quad \left. \left. \left. + y_{ies,m+1} \ln \left(1 - \sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m \frac{\exp(z'_{iesj} \beta + \kappa_l)}{1 + \sum_{\substack{k=1 \\ k \neq y_{ie0}}}^m \exp(z'_{iesk} \beta + \kappa_l)} \right) \right) \right] \right\} \\ &= \sum_{i=1}^N \ln \left[\sum_{l=1}^g \alpha_l \exp(u_l) \right] \end{aligned}$$

mit

$$\begin{aligned} \alpha_l &= \frac{w_l}{\sqrt{\pi}} \\ u_l &= \sum_{e=1}^{E_i} \sum_{s=1}^{t_e-1+\delta_{ie}} \left(\sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m y_{iesj} \ln e_{iesjl} + y_{ies,m+1} \ln \left(1 - \sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m e_{iesjl} \right) \right) \\ e_{iesjl} &= \frac{\exp(z'_{iesj} \beta + d_l \theta \sqrt{2})}{1 + \sum_{\substack{k=1 \\ k \neq y_{ie0}}}^m \exp(z'_{iesk} \beta + d_l \theta \sqrt{2})} \end{aligned}$$

$$S(\beta) = \begin{pmatrix} \frac{\partial \ln L}{\partial \beta^1} \\ \vdots \\ \frac{\partial \ln L}{\partial \beta^p} \\ \frac{\partial \ln L}{\partial \theta} \end{pmatrix}$$

im weiteren: $a, b = 1 \dots, |\beta| = p$

$$\begin{aligned} \frac{\partial \ln L}{\partial \beta^a} &= \sum_{i=1}^N \frac{1}{\sum_{l=1}^g \alpha_l \exp(u_l)} \cdot \left(\sum_{l=1}^g \alpha_l \exp(u_l) \cdot \frac{\partial u_l}{\partial \beta^a} \right) \\ &= \sum_{i=1}^N \frac{1}{\sum_{l=1}^g \alpha_l \exp(u_l)} \cdot \left(\sum_{l=1}^g \alpha_l \exp(u_l) \cdot v_{la} \right) \\ \frac{\partial \ln L}{\partial \theta} &= \sum_{i=1}^N \frac{1}{\sum_{l=1}^g \alpha_l \exp(u_l)} \cdot \left(\sum_{l=1}^g \alpha_l \exp(u_l) \cdot \frac{\partial u_l}{\partial \theta} \right) \\ &= \sum_{i=1}^N \frac{1}{\sum_{l=1}^g \alpha_l \exp(u_l)} \cdot \left(\sum_{l=1}^g \alpha_l \exp(u_l) \cdot v_{ld} \right) \end{aligned}$$

mit

$$\begin{aligned} \frac{\partial u_l}{\partial \beta^a} = v_{la} &= \sum_{e=1}^{E_i} \sum_{s=1}^{t_e-1+\delta_{ie}} \left(\sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m z_{iesj}^a (y_{iesj} - e_{iesjl}) \right) \\ \frac{\partial u_l}{\partial \beta^b} = v_{lb} &= \sum_{e=1}^{E_i} \sum_{s=1}^{t_e-1+\delta_{ie}} \left(\sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m z_{iesj}^b (y_{iesj} - e_{iesjl}) \right) \\ \frac{\partial u_l}{\partial \theta} = v_{ld} &= \sum_{e=1}^{E_i} \sum_{s=1}^{t_e-1+\delta_{ie}} \left(\sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m d_l \sqrt{2} (y_{iesj} - e_{iesjl}) \right) \end{aligned}$$

Zur Berechnung der Hesse-Matrix:

$$\begin{aligned} \frac{\partial^2 \ln L}{\partial \beta^a \partial \beta^b} &= \sum_{i=1}^N \frac{\partial^2 \ln L_i}{\partial \beta^a \partial \beta^b} \\ &= \sum_{i=1}^N \left\{ \frac{\left[\frac{\partial}{\partial \beta^b} \sum_{l=1}^g \alpha_l \exp(u_l) v_{la} \right] \left[\frac{\partial}{\partial \beta^a} \left(\sum_{l=1}^g \alpha_l \exp(u_l) \right) \right] \left[\sum_{l=1}^g \alpha_l \exp(u_l) v_{la} \right]}{\sum_{l=1}^g \alpha_l \exp(u_l) \left[\sum_{l=1}^g \alpha_l \exp(u_l) \right]^2} \right\} \end{aligned}$$

mit

$$\begin{aligned} \frac{\partial}{\partial \beta^b} \left(\sum_{l=1}^g \alpha_l \exp(u_l) v_{la} \right) &= \sum_{l=1}^g \left\{ a_l \exp(u_l) v_{lb} v_{la} + a_l \exp(u_l) \right. \\ &\quad \cdot \left[\sum_{e=1}^{E_i} \sum_{s=1}^{t_e - 1 + \delta_{ie}} \sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m \left(-z_{iesj}^a z_{iesj}^b e_{iesj} \right) \right. \\ &\quad \left. \left. \cdot \left(1 + \sum_{\substack{k=1 \\ k \neq y_{ie0}}}^m \exp(z'_{iesk} \beta + d_l \theta \sqrt{2}) \right)^{-1} \right] \right\} \\ \frac{\partial}{\partial \beta^b} \left(\sum_{l=1}^g a_l \exp(u_l) \right) &= \sum_{l=1}^g a_l \exp(u_l) v_{lb} \end{aligned}$$

ergibt sich

$$\begin{aligned} \frac{\partial^2 \ln L_i}{\partial \beta^a \partial \beta^b} &= \sum_{l=1}^g a_l \exp(u_l) \\ &\quad \frac{\left[v_{lb} v_{la} - \sum_{e=1}^{E_i} \sum_{s=1}^{t_e - 1 + \delta_{ie}} \sum_{\substack{j=1 \\ j \neq y_{ie0}}}^m z_{iesj}^a z_{iesj}^b e_{iesj} \left(1 + \sum_{\substack{k=1 \\ k \neq y_{ie0}}}^m \exp(z'_{iesk} \beta + d_l \theta \sqrt{2}) \right)^{-1} \right]}{\sum_{l=1}^g a_l \exp(u_l)} \\ &\quad \frac{\left[\sum_{l=1}^g a_l \exp(u_l) v_{lb} \right] \left[\sum_{l=1}^g a_l \exp(u_l) v_{la} \right]}{\left[\sum_{l=1}^g a_l \exp(u_l) \right]^2} \end{aligned}$$

Für die weiteren Ableitungen müssen entsprechend

$$\begin{aligned} \text{bei } \frac{\partial^2 \ln L_i}{\partial \beta^a \partial \theta} & \quad z_{isj}^b \text{ durch } d_l \sqrt{2} \text{ und } v_{lb} \text{ durch } v_{ld} \text{ ersetzt werden,} \\ \text{bei } \frac{\partial^2 \ln L_i}{\partial \theta \partial \beta^b} & \quad z_{isj}^a \text{ durch } d_l \sqrt{2} \text{ und } v_{la} \text{ durch } v_{ld} \text{ ersetzt werden,} \\ \text{bei } \frac{\partial^2 \ln L_i}{\partial \theta^2} & \quad z_{isj}^a \text{ und } z_{isj}^b \text{ jeweils durch } d_l \sqrt{2} \text{ und} \\ & \quad v_{la} \text{ und } v_{lb} \text{ jeweils durch } v_{ld} \text{ ersetzt werden.} \end{aligned}$$

B Auswertung der Gliomstudie

In den Tabellen 1 und 2 werden die Kovariablen epileptische Anfälle, Cortison und organisches Psychosyndrom zeitkonstant, in Tabelle 3 und 4 zeitveränderlich ausgewertet. Dabei stellt jeweils Tabelle 1 und 3 eine Analyse ohne Berücksichtigung und Tabelle 2 und 4 eine Analyse mit Berücksichtigung unbeobachteter Heterogenität dar.

Tabelle 5 und 6 enthalten die Auswertungsergebnisse mit zeitveränderlichen Kovariablen und einer Modellierung der Baselineparameter mit Hilfe der Verweildauer. In Tabelle 5 ist die Verweildauer die Aufenthaltsdauer der jeweiligen Untersuchungseinheit in der aktuellen Episode, in Tabelle 6 die Zeit, die seit Studieneintritt vergangen ist. Die Tabellen 5, 6 und 7 sind für mit und ohne Berücksichtigung unbeobachteter Heterogenität identisch, da die Standardabweichung der unbeobachteten Heterogenität fast gleich Null geschätzt wird. Ein Beispiel für eine Analyse, wenn der Zustand der vergangenen Episode mitgeschätzt wird, ist in Tabelle 7 angegeben. Die Auswertung erfolgt hier über eine Modellierung der Baselineparameter mittels der Verweildauer in der Studie.

Tabelle 1: Ohne unbeobachtete Heterogenität, zeitkonstante Kovariablen

	Tod	<71%	71-80%	81-100%
Baseline	-2.9879	-2.7187	-2.3671	-3.2755
(Varianz)	0.0279	0.0400	0.0285	0.0497
(p-value)	0.0000	0.0000	0.0000	0.0000
Therapie	-0.1434	-0.1558	0.1454	-0.0659
(Varianz)	0.0132	0.0198	0.0151	0.0251
(p-value)	0.2112	0.2683	0.2369	0.6776
Biopsie	0.7806	0.3944	-0.0071	-0.5973
(Varianz)	0.0441	0.0893	0.0820	0.2680
(p-value)	0.0002	0.1869	0.9804	0.2486
Totalresektion	-0.1056	-0.3867	-0.2616	0.4421
(Varianz)	0.0156	0.0254	0.0166	0.0260
(p-value)	0.3982	0.0152	0.0424	0.0061
Malign1	-0.5690	-0.3167	-0.3858	-0.2957
(Varianz)	0.0136	0.0152	0.0115	0.0179
(p-value)	0.0000	0.0103	0.0003	0.0272
Malign2	0.4265	0.1587	0.1697	0.1144
(Varianz)	0.0080	0.0106	0.0078	0.0127
(p-value)	0.0000	0.1237	0.0553	0.3093
Alter1	-0.3966	-0.4967	-0.0944	0.1191
(Varianz)	0.0068	0.0103	0.0074	0.0130
(p-value)	0.0000	0.0000	0.2740	0.2957
Alter2	0.0525	0.1182	0.0198	0.0194
(Varianz)	0.0063	0.0095	0.0079	0.0146
(p-value)	0.5070	0.2255	0.8241	0.8724
Bewusst.stör.	0.5394	0.8743	0.2033	-0.3104
(Varianz)	0.1605	0.2446	0.2346	0.6514
(p-value)	0.1781	0.0771	0.6747	0.7006
Epilept. Anfälle	-0.2881	-0.0781	-0.4685	0.2763
(Varianz)	0.0447	0.0599	0.0489	0.0636
(p-value)	0.1730	0.7496	0.0341	0.2733
Cortison	0.2804	0.0740	0.3541	-0.0963
(Varianz)	0.0158	0.0234	0.0173	0.0266
(p-value)	0.0258	0.6289	0.0071	0.5550
Organ.Psychosy.	-0.0210	0.1151	-0.2044	-0.0486
(Varianz)	0.0136	0.0205	0.0154	0.0256
(p-value)	0.8569	0.4212	0.0997	0.7612

Tabelle 2: Mit unbeobachteter Heterogenität, zeitkonstante Kovariablen

	Tod	<71%	71-80%	81-100%
Baseline	-2.9644	-2.7075	-2.3403	-3.2657
(Varianz)	0.0303	0.0424	0.0316	0.0526
(p-value)	0.0000	0.0000	0.0000	0.0000
Therapie	-0.1304	-0.1320	0.1489	-0.0333
(Varianz)	0.0148	0.0217	0.0172	0.0274
(p-value)	0.2841	0.3708	0.2559	0.8408
Biopsie	0.7789	0.3901	-0.0096	-0.5995
(Varianz)	0.0502	0.0975	0.0928	0.2808
(p-value)	0.0005	0.2115	0.9748	0.2579
Totalresektion	-0.1016	-0.3883	-0.2494	0.4428
(Varianz)	0.0175	0.0275	0.0189	0.0283
(p-value)	0.4423	0.0191	0.0697	0.0085
Malign1	-0.5775	-0.3175	-0.3764	-0.3236
(Varianz)	0.0150	0.0170	0.0132	0.0198
(p-value)	0.0000	0.0149	0.0011	0.0215
Malign2	0.4177	0.1564	0.1488	0.0958
(Varianz)	0.0089	0.0117	0.0090	0.0139
(p-value)	0.0000	0.1480	0.1165	0.4166
Alter1	-0.4001	-0.5037	-0.0972	0.1146
(Varianz)	0.0076	0.0113	0.0084	0.0141
(p-value)	0.0000	0.0000	0.2904	0.3345
Alter2	0.0577	0.1411	0.0088	0.0256
(Varianz)	0.0071	0.0106	0.0090	0.0158
(p-value)	0.4930	0.1695	0.9256	0.8385
Bewust.stör.	0.5380	0.9188	0.1787	-0.2820
(Varianz)	0.1837	0.2831	0.2590	0.6876
(p-value)	0.2094	0.0842	0.7256	0.7337
Epilept. Anfälle	-0.3048	-0.0462	-0.4939	0.2280
(Varianz)	0.0498	0.0661	0.0545	0.0701
(p-value)	0.1722	0.8575	0.0344	0.3894
Cortison	0.2736	0.0572	0.3681	-0.0778
(Varianz)	0.0177	0.0255	0.0196	0.0291
(p-value)	0.0395	0.7202	0.0086	0.6488
Organ.Psychosy.	-0.0158	0.1333	-0.1953	-0.0645
(Varianz)	0.0153	0.0224	0.0175	0.0278
(p-value)	0.8982	0.3736	0.1396	0.6991
Std. d. unbeob. H.	0.3592			
(Varianz)	0.0030			

Tabelle 3: Ohne unbeobachtete Heterogenität, zeitvariierende Kovariablen

	Tod	<71%	71-80%	81-100%
Baseline	-3.8634	-3.0733	-2.4662	-3.0266
(Varianz)	0.0350	0.0402	0.0284	0.0493
(p-value)	0.0000	0.0000	0.0000	0.0000
Therapie	-0.1084	-0.1582	0.1641	-0.0481
(Varianz)	0.0134	0.0198	0.0152	0.0253
(p-value)	0.3481	0.2611	0.1830	0.7622
Biopsie	0.7598	0.3939	0.0373	-0.7114
(Varianz)	0.0439	0.0891	0.0811	0.2665
(p-value)	0.0003	0.1869	0.8959	0.1682
Totalresektion	-0.0307	-0.3407	-0.2455	0.4180
(Varianz)	0.0156	0.0246	0.0167	0.0257
(p-value)	0.8058	0.0298	0.0571	0.0091
Malign1	-0.5697	-0.3233	-0.3528	-0.1976
(Varianz)	0.0140	0.0156	0.0116	0.0182
(p-value)	0.0000	0.0096	0.0011	0.1426
Malign2	0.3205	0.0982	0.1302	0.1193
(Varianz)	0.0083	0.0109	0.0081	0.0130
(p-value)	0.0004	0.3462	0.1493	0.2947
Alter1	-0.3475	-0.4388	-0.1060	0.1266
(Varianz)	0.0070	0.0106	0.0076	0.0132
(p-value)	0.0000	0.0000	0.2239	0.2710
Alter2	0.0519	0.1137	0.0010	-0.0095
(Varianz)	0.0064	0.0095	0.0080	0.0146
(p-value)	0.5155	0.2440	0.9910	0.9372
Bewust.stör.	0.2383	0.7469	0.1264	-0.1232
(Varianz)	0.1593	0.2392	0.2326	0.6622
(p-value)	0.5505	0.1267	0.7933	0.8796
Epil. zeitv.	0.4600	-0.1741	0.0165	-0.3298
(Varianz)	0.0169	0.0352	0.0255	0.0438
(p-value)	0.0004	0.3536	0.9174	0.1149
Cortison zeitv.	0.8364	0.2508	0.4726	0.3127
(Varianz)	0.0192	0.0229	0.0176	0.0283
(p-value)	0.0000	0.0977	0.0004	0.0631
Org. Psy. zeitv.	0.5151	0.5506	-0.2361	-0.7505
(Varianz)	0.0170	0.0229	0.0156	0.0258
(p-value)	0.0001	0.0003	0.0588	0.0000

Tabelle 4: Mit unbeobachteter Heterogenität, zeitvariierende Kovariablen

	Tod	<71%	71-80%	81-100%
Baseline	-3.8446	-3.0622	-2.4300	-3.0235
(Varianz)	0.0373	0.0425	0.0312	0.0514
(p-value)	0.0000	0.0000	0.0000	0.0000
Therapie	-0.0965	-0.1337	0.1656	-0.0212
(Varianz)	0.0148	0.0215	0.0169	0.0274
(p-value)	0.4275	0.3622	0.2030	0.8980
Biopsie	0.7592	0.3814	0.0399	-0.7129
(Varianz)	0.0491	0.0958	0.0898	0.2761
(p-value)	0.0006	0.2178	0.8940	0.1749
Totalresektion	-0.0307	-0.3462	-0.2312	0.4338
(Varianz)	0.0172	0.0265	0.0186	0.0279
(p-value)	0.8151	0.0335	0.0901	0.0094
Malign1	-0.5770	-0.3254	-0.3485	-0.2170
(Varianz)	0.0152	0.0171	0.0130	0.0199
(p-value)	0.0000	0.0129	0.0022	0.1241
Malign2	0.3138	0.1010	0.1172	0.1106
(Varianz)	0.0091	0.0118	0.0091	0.0141
(p-value)	0.0010	0.3523	0.2184	0.3517
Alter1	-0.3543	-0.4476	-0.1155	0.1255
(Varianz)	0.0077	0.0114	0.0084	0.0142
(p-value)	0.0001	0.0000	0.2080	0.2916
Alter2	0.0584	0.1341	-0.0052	-0.0034
(Varianz)	0.0071	0.0104	0.0088	0.0157
(p-value)	0.4870	0.1891	0.9559	0.9786
Bewust.stör.	0.2542	0.7972	0.1174	-0.1066
(Varianz)	0.1786	0.2717	0.2536	0.6957
(p-value)	0.5476	0.1261	0.8154	0.8983
Epil. zeitv.	0.4406	-0.1857	0.0090	-0.3541
(Varianz)	0.0180	0.0367	0.0270	0.0457
(p-value)	0.0010	0.3322	0.9563	0.0978
Cortison zeitv.	0.8320	0.2397	0.4700	0.3089
(Varianz)	0.0201	0.0240	0.0188	0.0296
(p-value)	0.0000	0.1215	0.0006	0.0725
Org. Psy. zeitv.	0.5199	0.5578	-0.2463	-0.7523
(Varianz)	0.0180	0.0241	0.0167	0.0272
(p-value)	0.0001	0.0003	0.0570	0.0000
Std. d. unbeob. H.	0.3287			
(Varianz)	0.0033			

Tabelle 5: Modellierung der Baseline: $\beta_{01j} + 0.1 \cdot s \cdot \beta_{02j}$, mit $s = 0, \dots, t_e$

	Tod	<71%	71-80%	81-100%
Baseline (konstant)	-4.1157	-2.2542	-1.7562	-2.0775
(Varianz)	0.0429	0.0465	0.0334	0.0551
(p-value)	0.0000	0.0000	0.0000	0.0000
Baseline (zeitabh.)	0.2178	-1.3876	-1.1160	-2.0014
(Varianz)	0.0038	0.0403	0.0248	0.0796
(p-value)	0.0004	0.0000	0.0000	0.0000
Therapie	-0.1258	-0.0788	0.1592	0.0763
(Varianz)	0.0133	0.0200	0.0155	0.0261
(p-value)	0.2758	0.5779	0.2007	0.6366
Biopsie	0.8056	0.2416	0.0001	-0.8545
(Varianz)	0.0439	0.0894	0.0833	0.2706
(p-value)	0.0001	0.4189	0.9999	0.1004
Totalresektion	-0.0410	-0.3247	-0.1350	0.4974
(Varianz)	0.0157	0.0247	0.0170	0.0264
(p-value)	0.7432	0.0387	0.3005	0.0022
Malign1	-0.6283	-0.1596	-0.1934	-0.0617
(Varianz)	0.0145	0.0161	0.0118	0.0188
(p-value)	0.0000	0.2092	0.0750	0.6526
Malign2	0.3599	0.0138	0.0299	-0.0239
(Varianz)	0.0086	0.0109	0.0082	0.0133
(p-value)	0.0001	0.8948	0.7414	0.8358
Alter1	-0.3728	-0.3433	-0.0534	0.2445
(Varianz)	0.0072	0.0106	0.0076	0.0134
(p-value)	0.0000	0.0008	0.5414	0.0347
Alter2	0.0498	0.1212	-0.0062	-0.0158
(Varianz)	0.0063	0.0097	0.0081	0.0150
(p-value)	0.5318	0.2188	0.9449	0.8976
Bewusst.stör.	0.2758	0.6847	-0.0101	-0.3489
(Varianz)	0.1586	0.2417	0.2327	0.6741
(p-value)	0.4885	0.1638	0.9833	0.6709
Epil.zeitv.	0.4635	-0.1135	0.0175	-0.2739
(Varianz)	0.0168	0.0360	0.0257	0.0454
(p-value)	0.0004	0.5498	0.9129	0.1987
Cortison zeitv.	0.9566	-0.0224	0.2515	0.0709
(Varianz)	0.0215	0.0229	0.0179	0.0289
(p-value)	0.0000	0.8821	0.0604	0.6765
Org.Psy.zeitv.	0.4914	0.5959	-0.2090	-0.6737
(Varianz)	0.0171	0.0233	0.0158	0.0262
(p-value)	0.0002	0.0001	0.0962	0.0000

Tabelle 6: Modellierung der Baseline: $\beta_{01j} + 0.1 \cdot s \cdot \beta_{02j}$, mit $s = 0, \dots, \sum_e t_e$

	Tod	<71%	71-80%	81-100%
Baseline (konstant)	-4.2246	-2.0240	-1.4406	-1.6517
(Varianz)	0.0462	0.0482	0.0354	0.0568
(p-value)	0.0000	0.0000	0.0000	0.0000
Baseline (zeitabh.)	0.2147	-1.1301	-1.0784	-1.7858
(Varianz)	0.0027	0.0172	0.0131	0.0348
(p-value)	0.0000	0.0000	0.0000	0.0000
Therapie	-0.1417	-0.0307	0.1860	0.1206
(Varianz)	0.0134	0.0203	0.0159	0.0269
(p-value)	0.2204	0.8292	0.1404	0.4620
Biopsie	0.8395	0.1383	-0.0819	-0.9861
(Varianz)	0.0439	0.0912	0.0857	0.2778
(p-value)	0.0001	0.6471	0.7798	0.0613
Totalresektion	-0.0472	-0.2906	-0.1199	0.5032
(Varianz)	0.0155	0.0251	0.0174	0.0273
(p-value)	0.7050	0.0667	0.3641	0.0023
Malign1	-0.6296	-0.1719	-0.2043	-0.0121
(Varianz)	0.0143	0.0164	0.0120	0.0195
(p-value)	0.0000	0.1796	0.0624	0.9312
Malign2	0.3769	-0.0343	-0.0126	-0.0776
(Varianz)	0.0087	0.0110	0.0083	0.0134
(p-value)	0.0001	0.7440	0.8902	0.5032
Alter1	-0.4066	-0.2455	0.0283	0.3615
(Varianz)	0.0075	0.0108	0.0079	0.0137
(p-value)	0.0000	0.0180	0.7493	0.0020
Alter2	0.0537	0.0986	-0.0053	-0.0465
(Varianz)	0.0063	0.0098	0.0082	0.0153
(p-value)	0.4997	0.3202	0.9538	0.7069
Bewust.stör.	0.2312	0.7693	-0.0839	-0.5692
(Varianz)	0.1572	0.2467	0.2356	0.6989
(p-value)	0.5598	0.1214	0.8628	0.4960
Epil.zeitv.	0.4597	-0.0707	0.0656	-0.1428
(Varianz)	0.0168	0.0367	0.0264	0.0470
(p-value)	0.0004	0.7121	0.6862	0.5102
Cortison zeitv.	0.9874	-0.0677	0.1915	0.0456
(Varianz)	0.0218	0.0230	0.0180	0.0296
(p-value)	0.0000	0.6553	0.1572	0.7912
Org.Psy.zeitv.	0.4636	0.6845	-0.1281	-0.6036
(Varianz)	0.0170	0.0238	0.0162	0.0273
(p-value)	0.0004	0.0000	0.3142	0.0003

Tabelle 7: Modellierung der Baseline: $\beta_{01j} + 0.1 \cdot s \cdot \beta_{02j}$, mit $s = 0, \dots, \sum_e t_e$

	Tod	<71%	71-80%	81-100%
Baseline (konstant)	-4.1582	-1.9588	-1.5560	-1.6352
(Varianz)	0.0459	0.0486	0.0384	0.0597
(p-value)	0.0000	0.0000	0.0000	0.0000
Baseline (zeitabh.)	0.2070	-1.1769	-1.1713	-1.8692
(Varianz)	0.0027	0.0223	0.0162	0.0472
(p-value)	0.0001	0.0000	0.0000	0.0000
Zust. ehe.Episoden: <71%	0.3624	0.2799	0.2506	-0.6641
(Varianz)	0.0090	0.1227	0.0257	0.0871
(p-value)	0.0001	0.4241	0.1181	0.0244
Zust. ehe.Episoden: 71-80%	-0.0963	0.2368	-0.5593	0.3275
(Varianz)	0.0090	0.0297	0.0765	0.0398
(p-value)	0.3109	0.1692	0.0431	0.1005
Therapie	-0.1436	-0.0542	0.1781	0.1163
(Varianz)	0.0135	0.0204	0.0158	0.0271
(p-value)	0.2166	0.7044	0.1563	0.4798
Biopsie	0.8137	0.0984	-0.0437	-1.0209
(Varianz)	0.0438	0.0903	0.0855	0.2776
(p-value)	0.0001	0.7434	0.8813	0.0526
Totalresektion	-0.0321	-0.2441	-0.1254	0.4906
(Varianz)	0.0157	0.0253	0.0175	0.0276
(p-value)	0.7979	0.1251	0.3434	0.0031
Malign1	-0.6582	-0.1783	-0.1902	-0.0039
(Varianz)	0.0146	0.0165	0.0119	0.0196
(p-value)	0.0000	0.1648	0.0817	0.9780
Malign2	0.3921	-0.0339	-0.0064	-0.0829
(Varianz)	0.0088	0.0111	0.0082	0.0135
(p-value)	0.0000	0.7472	0.9435	0.4763
Alter1	-0.3778	-0.2308	0.0311	0.3442
(Varianz)	0.0076	0.0107	0.0079	0.0137
(p-value)	0.0000	0.0259	0.7264	0.0033
Alter2	0.0337	0.1065	-0.0215	-0.0375
(Varianz)	0.0064	0.0098	0.0082	0.0153
(p-value)	0.6732	0.2826	0.8123	0.7620
Bewusst.stör.	0.2669	0.7459	-0.0309	-0.5926
(Varianz)	0.1576	0.2495	0.2352	0.7053
(p-value)	0.5014	0.1354	0.9493	0.4805
Epil.zeitv.	0.4431	-0.1029	0.0672	-0.1858
(Varianz)	0.0169	0.0366	0.0261	0.0471
(p-value)	0.0006	0.5907	0.6778	0.3921
Cortison zeitv.	0.9341	-0.0998	0.1960	0.0433
(Varianz)	0.0217	0.0232	0.0182	0.0304
(p-value)	0.0000	0.5119	0.1469	0.8039
Org.Psy.zeitv.	0.3860	0.6445	-0.1267	-0.5366
(Varianz)	0.0176	0.0241	0.0167	0.0278
(p-value)	0.0036	0.0000	0.3268	0.0013

Literatur

- [1] **Allison**, P. D. (1982). Discrete-time methods for the analysis of event histories. *Sociological Methodology, Vol 13*, S. 61-98.
- [2] **Aydemir**, S., **Aydemir**, Ü. und **Dirschedl**, P. (1996). *Survivalanalysen mit Berücksichtigung der zeitlichen Kovariablenentwicklung in klinischen Studien*. Discussion paper 44, Sonderforschungsbereich 386 der Ludwigs-Maximilians-Universität München.
- [3] **Blossfeld**, H. P., **Hamerle**, A. und **Mayer**, K. U. (1989). *Event History Analysis: Statistical Theory and Applications in Social Sciences*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- [4] **Davies**, P. J. und **Rabinowitz**, P. (1975). *Numerical Integration*. Waltham, Massachusetts: Blaisdell.
- [5] **Davies**, P. J. und **Rabinowitz**, P. (1984). *Methods of Numerical Integration*. Orlando, F.: Academic Press.
- [6] **Dennis**, J. E. jr. und **Schnabel**, R. B. (1983). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, New Jersey: Prentice-Hall.
- [7] **Fahrmeir**, L. und **Tutz**, G. (1994). *Multivariate Statistical Modelling Based on Generalized Linear Models*. New York, Berlin: Springer Verlag.
- [8] **Hastie**, T. und **Tibshirani**, R. (1993). Varying-coefficient Models. *Journal of the Royal Statistical Society B 55*, S. 757-796.
- [9] **Kalbfleisch**, J. D. und **Prentice**, R. L. (1980). *The Statistical Analysis of Failure Time Data*. New York: John Wiley & Sons.
- [10] **Lancaster**, T. (1990). *The Econometric Analysis of Transition Data*. Cambridge: Cambridge University Press.
- [11] **Lawless**, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. New York: Wiley.
- [12] **Petersen**, T. (1991). The statistical analysis of event histories. *Sociological Methods & Research, Vol. 19 No. 3*, S. 270-323.
- [13] **Scheike**, T. H. und **Jensen**, T. K. (1997). A discrete survival model with random effects: An application to time to pregnancy. *Biometrics 53*, S. 349-360.

- [14] **Stahl**, E. (1996). *Diskrete Verweildauermodelle mit zufälligen Effekten*. Frankfurt/M, Lang-Verlag.
- [15] **Thompson**, W. A. jr. (1977). On the treatment of grouped observations in life studies. *Biometrics* 33, S. 463-470.
- [16] **Ulm**, K., **Schmoor**, C., **Sauerbrei**, W., **Kemmler**, G., **Aydemir**, Ü., **Müller**, B. und **Schumacher**, M. (1989). Strategien zur Auswertung einer Therapiestudie mit der Überlebenszeit als Zielkriterium. *Biometrie und Informatik in Medizin und Biologie*, 20 (4), S. 171-205.
- [17] **Wangler**, A. (1997). *Heterogenitätsprobleme in der Verlaufsdatenanalyse*. Frankfurt/M, Lang-Verlag.
- [18] **Yamaguchi**, K. (1993). Modeling time-varying effects of covariates in event-history analysis using statistics from the saturated hazard rate model. *Sociological Methodology*, Vol. 23, S. 279-317.