

Aus dem Institut für Medizinische Epidemiologie, Biometrie und Informatik
(Direktor: Prof. Dr. Johannes Haerting)

Regressionsmodelle für nichtstetige Zielgrößen und korrelierte Beobachtungen

H a b i l i t a t i o n

zur Erlangung des akademischen Grades

Dr. sc. hum., rer. medic. habil.

vorgelegt der Medizinischen Fakultät
der Martin-Luther-Universität Halle-Wittenberg

von Dr. sc. hum. Oliver Kuß
geboren am 16. Juli 1969 in Crailsheim

Gutachter:

1. (wird später eingetragen)
2. (wird später eingetragen)
3. (wird später eingetragen)

Referat und bibliographische Gestaltung

Eine Reihe von Studienanlagen in der Medizin führt zu korrelierten oder abhängigen Beobachtungen. Patienten werden z.B. mehrmals im Zeitverlauf, unter verschiedenen experimentellen Bedingungen oder in logischen Einheiten (Clustern) beobachtet. Die Standardmethoden der Statistik, die in der Regel eine Unabhängigkeit der Beobachtungen voraussetzen, sind dann nicht mehr gültig. Dies gilt auch für die Gruppe der Regressionsmodelle, deren Ziel es ist, den Einfluss eines oder mehrerer unabhängiger Merkmale auf ein abhängiges Merkmal (Zielgröße) zu beschreiben. Regressionsmodelle mit stetigen Zielgrößen sind auch beim Vorliegen von korrelierten Beobachtungen inzwischen beherrschbar, komplizierter wird die Situation bei nichtstetigen (binären, nominalen oder diskreten) Zielgrößen.

In dieser Arbeit wird zum einen gezeigt, dass Regressionsmodelle für nichtstetige Zielgrößen und korrelierte Beobachtungen in der Medizin immer wichtiger werden. Dieses folgt aus der Analyse der angewandten Methoden im medizinischen Schrifttum im Verlauf der Zeit und aus der Tatsache, dass immer mehr nichttechnische Übersichtsartikel zu diesen Methoden erscheinen. Zum anderen werden anhand von vier konkreten Anwendungen Regressionsmodelle für nichtstetige Zielgrößen und korrelierte Beobachtungen dargestellt. Betrachtet werden dabei binäre, nominale und diskrete Zielgrößen, außerdem variiert das Anwendungsgebiet von der Kardiochirurgie über die Onkologie und Versorgungsforschung zur Allergologie.

Konkret wird gezeigt, dass bei Meta-Analysen für seltene Ereignisse der asymptotische Mantel-Haenszel-Schätzer dem exakten überlegen ist und wie korrelierte Überlebenszeiten aus Tumorregistern ausgewertet werden können. Des Weiteren wird dargestellt, wie Regressionsmodelle für nominale Zielgrößen und korrelierte Beobachtungen in SAS ohne die Anwendung einer Poisson-Likelihood effizienter geschätzt werden können und wie mit Hilfe eines Mischungsmodelles Gruppen von Neurodermitisverläufen datenabhängig identifiziert und dargestellt werden können.

Kuß, Oliver: Regressionsmodelle für nichtstetige Zielgrößen und korrelierte Beobachtungen. Halle, Univ., Med. Fak., Habil., 87 Seiten, 2008

Inhaltsverzeichnis

Inhaltsverzeichnis	3
Verzeichnis der Abkürzungen und Symbole	5
1. Einleitung.....	6
1.1 Korrelierte Beobachtungen.....	6
1.2 Regressionsmodelle.....	8
1.3 Nichtstetige Zielgrößen	11
2. Regressionsmodelle für nichtstetige Zielgrößen und korrelierte Beobachtungen..	15
2.1 Marginale versus konditionale Modelle	16
3. Anwendungen von Regressionsmodellen für nichtstetige Zielgrößen und korrelierte Beobachtungen in der Medizin	20
3.1 Ein exakter Test für die Meta-Analyse mit binären Endpunkten.....	22
3.2 Ein Regressionsmodell für die relative Überlebenszeit bei korrelierten Beobachtungen	24
3.3 Die Schätzung von nominalen logistischen Modellen mit korrelierten Beobachtungen in SAS	26
3.4 Datenabhängiges, modellbasiertes Gruppieren von binären longitudinalen Verläufen am Beispiel der Neurodermitis	29
4. Zusammenfassung.....	33
5. Literaturverzeichnis	34
6. Anlagen	42
Kuss O, Gromann C. An exact test for meta-analysis with binary endpoints. Methods of Information in Medicine 2007;46(6):662-8.	42
Kuss O, Blankenburg T, Haerting J. A relative survival model for clustered responses. Biometrical Journal 2008;50(3):408-18.	50
Kuss O, McLerran D. A note on the estimation of the multinomial logistic model with correlated responses in SAS. Computer Methods and Programs in Biomedicine 2007;87(3):262-9.	62

Kuss O, Gromann C, Diepgen TL. Model-based clustering of binary longitudinal atopic dermatitis disease histories by latent class mixture models. Biometrical Journal 2006;48(1):105-16..... 71

7. Thesen	84
8. Tabellarischer Lebenslauf	86

Verzeichnis der Abkürzungen und Symbole

ANOVA	Analysis of Variance, Varianzanalyse
BIC	Bayesian Information Criterion
DMW	Deutsche Medizinische Wochenschrift
GEE	Generalized Estimation Equations
GENMOD	Generalized Linear Models
GLIMMIX	Generalized Linear Models, Mixed
GLM	Generalized Linear Models
GLMM	Generalized Linear Mixed Models
HALLUCA	Halle Lung Carcinoma
ICC	Intra-Class-Korrelation
MAS	Multicenter Allergie-Studie
MCMC	Markov Chain Monte Carlo
MC	Multiple Choice
MKG	Musterkrankengeschichte
NEJM	New England Journal of Medicine
NLMIXED	Nonlinear Mixed
OR	Odds Ratio
PQL	Penalized Quasi-Likelihood
SAS	Statistical Analysis System
SHT	Schädel-Hirn-Trauma
SPSS	Statistical Package for the Social Sciences

1. Einleitung

1.1 Korrelierte Beobachtungen

Eine Reihe von Studienanlagen in der Medizin führt zu korrelierten oder abhängigen Beobachtungen:

- Patienten werden im Zeitverlauf beobachtet und Messungen werden an mehreren Zeitpunkten gemacht (z.B. jährliche Erhebung des Vorliegens von Neurodermitis-symptomen beim Heranwachsen eines Kindes¹).
- Patienten werden simultan unter verschiedenen experimentellen Bedingungen behandelt (z.B. Messung des transepidermalen Wasserverlustes nach Irritation gleichzeitig an verschiedenen Körperstellen²).
- Es werden mehrere Zielgrößen pro Patient beobachtet (z.B. Durchführung verschiedener kognitiver Tests bei Rehabilitanden mit psychischen Krankheiten³).
- Patienten werden in logischen Einheiten (Clustern) beobachtet, (z.B. als Patienten von verschiedenen Allgemeinpraxen, die bei Vorliegen einer Verdachtsdiagnose an eine spezialisierte Klinik überwiesen werden⁴).
- Im Rahmen eines systematischen Reviews wird eine Meta-Analyse von einzelnen Studien zur selben Fragestellung durchgeführt⁵. Diesbezüglich sind die Patienten innerhalb einer Studie als korreliert anzusehen.

Häufig findet man sogar mehrere dieser Phänomene gleichzeitig und sich überlagernd in einer Studie, also z.B. den Vergleich von Behandlungen an verschiedenen Zeitpunkten in verschiedenen Clustern⁶ oder die Erhebung von mehreren Zielgrößen im Zeitverlauf³.

Die Standardannahme in den meisten, grundlegenden statistischen Auswertungsverfahren (z.B. Chi-Quadrat-Test, lineare Regression) ist jedoch die Unabhängigkeit der Beobachtungen und eine Analyse von korrelierten Messungen mit den Standardmethoden für unkorrelierte Beobachtungen führt unter Umständen zu falschen Schlüssen. Folgendes einfaches Beispiel (vgl. ⁷) soll dies verdeutlichen.

Wir nehmen an, mit einer Gruppe von 50 adipösen Kindern wird ein 4-wöchiges Sportprogramm zur Gewichtsreduktion durchgeführt, das Gewicht der Kinder wird vor Beginn und zum Ende des Sportprogrammes gemessen. Das Ausgangsgewicht der Kinder soll normalverteilt sein mit Erwartungswert 80 kg und Varianz $\sigma^2 = 25 (= 5^2)$, so dass gemäß der 2σ -Regel ca. 95% der Kinder ein Ausgangsgewicht zwischen 70 und 90 kg haben. Als Ergebnis der Studie zeigt sich, dass die Kinder im Mittel 2 kg abgenommen haben, genauer, soll die Gewichtsabnahme normalverteilt mit Erwartungswert 2 kg und Varianz 1 sein, die Gewichtsabnahme wird dabei als unabhängig vom Ausgangsgewicht angenommen. Es ist klar, dass die methodisch saubere Evaluierung eines Sportprogramms eine Kontrollgruppe mitführen müsste, vorzugsweise zusätzlich eine randomisierte Zuteilung der Kinder zu den Gruppen, aber wir vernachlässigen diesen Einwand für den Moment.

Führen wir eine Analyse durch, die *nicht* berücksichtigt, dass die Messwerte der beiden Gruppen ("Gewicht in kg *vor* dem Sportprogramm" und "Gewicht in kg *nach* dem Sportprogramm") korreliert sind, da jedes Kind Gewichtswerte in beiden Gruppen beisteuert, dann ist ein unverbundener t-Test (oder: t-Test für unabhängige Stichproben) die nahe liegende Auswertungsmethode. Die Formel zur Berechnung der Test-Statistik des t-Tests t_{uv} sieht dann, etwas vereinfacht, so aus:

$$t_{uv} = \frac{\text{Differenz der Mittelwerte}}{\text{Mittlere Standardabweichung}} * f_{uv}(N)$$

d.h. die Differenz der Mittelwerte in beiden Gruppen, als naheliegender Schätzer für den Effekt des Sportprogramms, wird standardisiert an einer mittleren Standardabweichung der Gewichtswerte in beiden Gruppen, zusätzlich ist der Wert der Test-Statistik noch eine Funktion ($f_{uv}(N)$) der Fallzahl. In unserem konkreten Beispiel ergibt sich als Wert der Teststatistik $t_{uv} = 2,00$, was bei einer Anzahl von 98 Freiheitsgraden einem p-Wert von $p = 0,048$ entspricht.

Eine korrekte Analyse der vorliegenden Daten, die berücksichtigt, dass eigentlich korrelierte Messungen vorliegen, würde naheliegenderweise einen t-Test für verbundene (oder: abhängige) Stichproben heranziehen. Die Test-Statistik des t-Tests für verbundene Stichproben t_v sieht dann, ebenfalls etwas vereinfacht, so aus:

$$t_v = \frac{\text{Mittelwert der Differenzen}}{\text{Standardabweichung der Differenzen}} * f_v(N).$$

Der Effekt des Sportprogramms wird nicht mehr als Differenz der Mittelwerte in der "Vorher"- und "Nachher"-Gruppe gemessen, sondern jetzt wird, für jedes Kind einzeln, zuerst eine Differenz zwischen "Vorher"- und "Nachher"-Wert berechnet. Der Mittelwert über diese Differenzen stellt dann einen Schätzer für den Effekt des Sportprogramms dar. Noch entscheidender in unserem Beispiel aber ist der Nenner der Test-Statistik. Hier wird jetzt an der Standardabweichung der gemessenen Differenzen standardisiert. Diese ist in unserem Beispiel viel kleiner als im unverbundenen Fall, weil die Streuung der Ausgangsgewichte nicht mit berücksichtigt werden muss. In unserem Beispiel ergibt sich ein Wert der Teststatistik von $t_v = 14,24$, was bei einer Anzahl von 49 Freiheitsgraden zu einem höchst-signifikanten p-Wert führt (der, um ganz genau zu sein, gleich $4,7 \cdot 10^{-19}$ ist.).

Zugegebenermaßen handelt es sich hier um ein etwas konstruiertes Beispiel, aber ein Prinzip der statistischen Verfahren für korrelierte Messungen wird klar: Jeder Patient ist seine eigene Kontrolle und die zusätzliche Variabilität, die unabhängige Kontrollen mit sich bringen würden, muss nicht mehr berücksichtigt werden.

1.2 *Regressionsmodelle*

Alle in der Einleitung dargestellten medizinischen Probleme (und noch viele mehr in der medizinischen Forschung) lassen sich aus statistischer Sicht in der Form eines Regressionsmodells darstellen. Ein Regressionsmodell ist immer dann gegeben, wenn wir uns für den Einfluss eines (X) oder mehrerer Merkmale (X_1, X_2, \dots, X_m , auch: Kovariablen, unabhängige Variablen, Prädiktoren) auf eine Zielgröße Y (auch: abhängige Variable) interessieren. In seiner einfachsten Form besteht ein Regressionsmodell nur aus einer einzelnen Kovariable X und einer Zielgröße Y . Sind beide dieser Merkmale zudem noch als stetige Merkmale anzusehen, ergibt sich ein lineares Regressionsmodell in seiner einfachsten Form als Geradengleichung¹

¹ Im Prinzip müsste diese Geradengleichung auch noch einen Term ε für den Messfehler enthalten. Dieser wird aus didaktischen Gründen ausgespart.

$$Y = \alpha + \beta X.$$

In diesem simplen Fall ist das vorliegende Problem auch noch einer eingängigen grafischen Darstellung, dem so genannten Streudiagramm (oder Scatterplot) zugänglich. In Abbildung 1 ist der Zusammenhang zwischen der unabhängigen Variable X „Körpergröße (in cm)“ und der abhängigen Variable Y „Körpergewicht (in kg)“ in einer Gruppe von 604 Teilnehmern einer randomisierten Studie zum Vergleich zweier Operationstechniken (mit oder ohne Herz-Lungen-Maschine) in der Kardiochirurgie⁸ dargestellt. Ein Anwachsen der Größe mit zunehmendem Gewicht ist offensichtlich und wird durch die eingezeichnete Regressionsgerade beschrieben. Die Steigung β der Regressionsgerade beträgt hier 0,82 [95%-Konfidenzintervall: 0,71; 0,93] und kann folgendermaßen interpretiert werden: Mit jedem Zentimeter Zuwachs an Körpergröße ergibt sich ein mittlerer Zuwachs von 0,82 kg Körpergewicht. Dabei ist jedoch die unterstellte Linearität zu beachten. Es wird bei der Berechnung explizit angenommen, dass dieser Zusammenhang für alle Körpergrößen im beobachteten Bereich gilt. Weiterhin wesentlich ist die Abwesenheit eines echt kausalen Zusammenhangs zwischen Körpergröße und Körpergewicht. Die hier unterstellte Beziehung (mit dem Zuwachs an Körpergröße steigt auch das Gewicht) bedeutet nicht, dass der Gewichtszuwachs ursächlich auf das Anwachsen der Körpergröße zurückzuführen ist.

Es ist physiologisch offensichtlich, dass ein Zusammenhang nicht nur zwischen der Körpergröße und dem Körpergewicht besteht, sondern auch zwischen dem Alter und dem Körpergewicht. Der gemeinsame Einfluss (in einem nichtkausalen Sinne) von Alter und Körpergröße auf das Körpergewicht kann also in einem multiplen Regressionsmodell

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2$$

dargestellt werden, wobei X_1 für das Merkmal „Körpergröße (in cm)“ steht und X_2 für das Merkmal „Alter (in Jahren)“. Analog zum vorigen Fall einer einzelnen unabhängigen Variablen können auch hier wieder die zugehörigen Regressionsparameter β_1 und β_2 bestimmt und als Steigungsmaße interpretiert werden. Es ergeben sich hier für die Körpergröße 0,80 [0,68; 0,91] und für das Alter -0,13 [-0,24; -0,02]. Der große

Vorteil eines multiplen Regressionsmodells ist, dass der Einfluss eines jeden unabhängigen Merkmales auf die Zielgröße um den Einfluss aller jeweils anderen unabhängigen Merkmale im Modell bereinigt ist. Das heißt, bereinigt für das zugrundeliegende Alter der Patienten, ergibt sich mit jedem Zentimeter Zuwachs an Körpergröße ein mittlerer Zuwachs von 0,80 kg Körpergewicht (und nicht mehr 0,82 wie im unadjustierten Fall). Analoges gilt für das Merkmal Alter: Bereinigt für die zugrundeliegende Körpergröße der Patienten, ergibt sich mit jedem Lebensjahr eine Gewichtsabnahme von 0,13 kg. Zum Vergleich, der unadjustierte Regressionskoeffizient für das Alter ist $-0,25$ $[-0,37; -0,12]$. Ein beträchtlicher Teil des Einfluss des Alters auf das Körpergewicht kann also durch den Einfluss der Körpergröße erklärt werden.

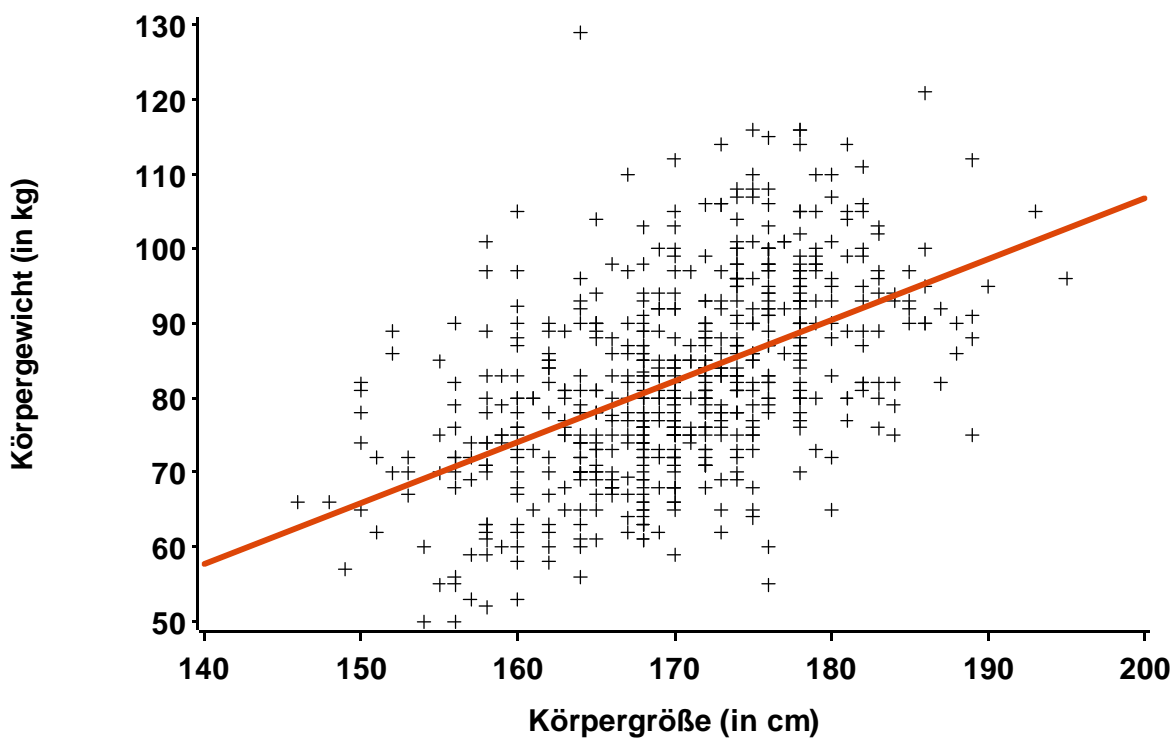


Abbildung 1: Zusammenhang zwischen der unabhängigen Variable X „Körpergröße (in cm)“ und der abhängigen Variable Y „Körpergewicht (in kg)“ in einer Gruppe von 604 Teilnehmern einer randomisierten Studie zum Vergleich zweier Operationsmethoden (mit oder ohne Herz-Lungen-Maschine) in der Kardiochirurgie

1.3 Nichtstetige Zielgrößen

Nicht immer (oder besser: eher selten) hat man es in der Medizin mit stetigen Zielgrößen zu tun. Viel häufiger treten nichtstetige Zielgrößen auf, unter denen die binären Zielgrößen (z.B. das Vorliegen einer Krankheit oder das Eintreten eines unerwünschten Ereignisses) die prominenteste Gruppe darstellt. Aber auch kategorielle Zielgrößen mit mehr als zwei Kategorien treten auf, unter diesen werden solche mit nominalem (ohne explizite inhaltliche Ordnung der Kategorien) oder ordinalem (mit expliziter inhaltlichen Ordnung der Kategorien) Skalenniveau unterschieden. Das wesentliche Unterscheidungsmerkmal der verschiedenen Skalenniveaus ist dabei (vgl. Abbildung 2) die Anzahl der Ausprägungen eines Merkmales. Selbst unter den Merkmalen mit unendlich vielen Ausprägungen müssen noch einmal solche mit abzählbar oder überabzählbar unendlich vielen Ausprägungen unterschieden werden. Diese Differenzierung ist vor allem aus lokaler Sicht bemerkenswert, da die Unterscheidung zwischen diesen beiden Begriffen der Unendlichkeit von einem halleschen Mathematiker, dem Begründer der mathematischen Mengenlehre, Georg Cantor (1845 - 1918) eingeführt worden war. Überdies hat die Zeit von Georg Cantor als Mathematik-Professor in Halle auch eine starke medizinische Komponente. Georg Cantor musste aufgrund einer psychischen Erkrankung viele Jahre in der damaligen Irrenanstalt, der heutigen Universitätsklinik und Poliklinik für Psychiatrie, Psychotherapie und Psychosomatik der medizinischen Fakultät verbringen.

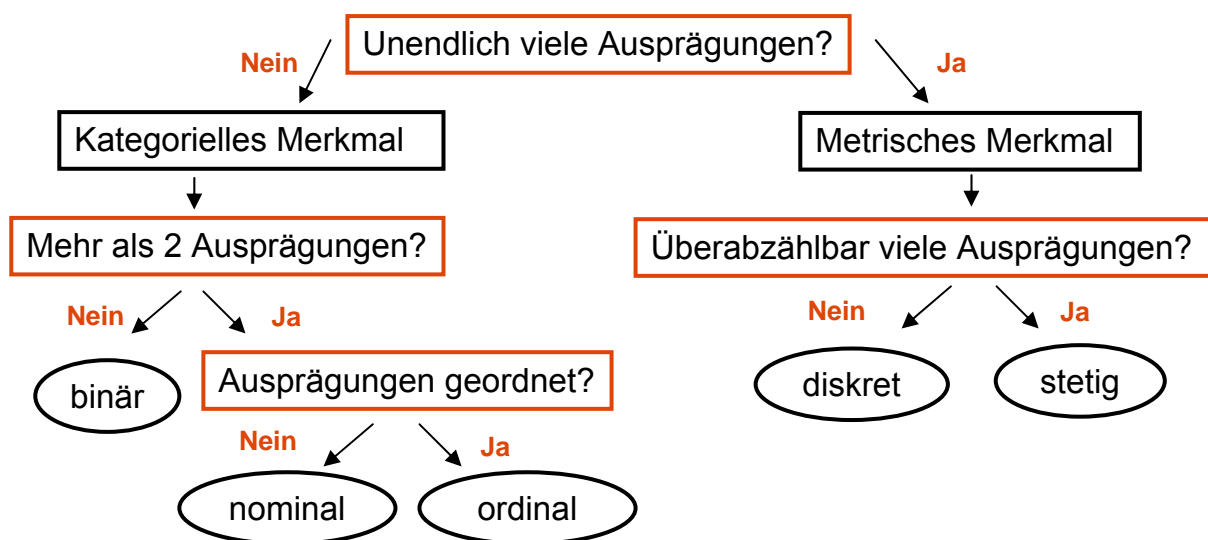


Abbildung 2: Darstellung und Unterschiede verschiedener Skalenniveaus

Eine sorgfältige Bestimmung des Skalenniveaus der Zielgröße ist in der medizinischen Statistik deshalb wichtig, weil sich daraus das zu verwendende Regressionsmodell ableitet. Wie wir bereits gesehen haben, ist für eine metrische Zielgröße das klassische lineare Regressionsmodell anzuwenden. Für binäre Zielgrößen findet im allgemeinen das logistische Regressionsmodell Anwendung, für nominale das multinomiale logistische Modell, für ordinale das Proportional Odds Modell⁹ und für diskrete das Poisson-Modell. Lineares, logistisches und Poisson-Modell können als Spezialfälle der Familie der Generalisierten Linearen Modelle (GLM)¹⁰ angesehen werden. Ein von dieser Klassifizierung abweichendes, aber in der Medizin sehr häufig angewandtes Regressionsmodell ist das Proportional Hazards-Modell (oder Cox-Modell)¹¹ für zensierte Überlebenszeiten. Unter Zensierung versteht man hier die Tatsache, dass das Zielereignis (z.B. Tod) bis zum Ende der Studie noch nicht beobachtet worden ist. Eine Bestimmung der Zeit bis zum Eintreten des Ereignisses (= Überlebenszeit) ist im Falle einer zensierten Beobachtung nicht möglich, eine Gleichsetzung der Beobachtungszeit mit der Überlebenszeit würde zu einer Unterschätzung der Überlebenszeit führen. Im Kontext der Regressionsmodelle würde eine Anwendung des linearen Regressionsmodells (mit dem Argument, bei der Überlebenszeit handele es sich um ein stetiges Merkmal) in dieser Situation zu falschen Ergebnissen führen, da das lineare Regressionsmodell Zensierungen nicht mit berücksichtigen kann. Es ist zu beachten, dass ausschließlich das Skalenniveau *der Zielgröße* die Auswahl des anzuwendenden Regressionsmodells bestimmt. Unabhängig davon sind die Skalenniveaus der *unabhängigen Variablen*, alle Regressionsmodelle können sowohl kategorielle als auch metrische Merkmale als Kovariablen modellieren.

1.3.1 Regressionsmodelle für nichtstetige Zielgrößen

Das einfachste Regressionsmodell für nichtstetige Zielgrößen ist ein logistisches Regressionsmodell mit einer einzelnen binären Kovariablen. Eine intuitive graphische Darstellung wie beim linearen Regressionsmodell ist hier nicht mehr ohne weiteres möglich, allerdings kann die wohlbekanntere Vierfeldertafel herangezogen werden, um die Daten darzustellen. Als Beispiel dient hier eine multizentrische randomisierte Studie zum Vergleich zweier topischer Cremes (Verum vs. Kontrolle) zur Behandlung nichtspezifischer Infektionen. Diese Studie war erstmals von Beitler/Landis¹² in einer methodischen Arbeit als Anwendungsbeispiel herangezogen worden und wurde auch

in der Folgezeit häufig benutzt, um Methoden der Regressionsanalyse für korrelierte Beobachtungen mit binären Zielgrößen darzustellen¹³⁻¹⁷.

Folgende Vierfeldertafel zeigt die Ergebnisse der Studie:

		Heilung		
		Nein	Ja	
Behandlung	Kontrolle	96	47	143
	Verum	75	55	130
		171	102	273

Eine Reihe von Effektmaßen könnte herangezogen werden, um den Behandlungseffekt zu messen, wir verwenden hier, primär aus statistischen Gründen, das Odds Ratio. Es ergibt sich ein Odds Ratio von 1,50 [0,92; 2,45], der p-Wert des zugehörigen χ^2 -Tests ist $p = 0,11$. Das heißt, die Chance unter dem Verum geheilt zu werden ist um 50% erhöht im Vergleich zur Kontrolle, wobei dieser Effekt allerdings nicht statistisch signifikant ist.

Wie bereits angedeutet, kann die zugrundeliegende Fragestellung auch mit einem logistischen Regressionsmodell modelliert werden, wobei die Heilung die binäre Zielgröße darstellt und die Behandlung die einzige binäre Kovariable. Konkret ergibt sich die Regressionsgleichung

$$\log(\pi / (1-\pi)) = \alpha + \beta X$$

mit $\pi = P(Y = 1)$, wobei Y für die binäre Zielgröße steht. Die rechte Seite der Modellgleichung ist gleich der im linearen Regressionsmodell, d.h. auch hier wird wieder ein linearer Einfluss der Kovariablen angenommen. Unterschiedlich ist allerdings die linke Seite der Gleichung. Es wird nicht direkt die Zielgröße Y modelliert, sondern die Wahrscheinlichkeit π , dass die Zielgröße eintritt, im o.g. Beispiel also, dass ein Patient von seiner Infektion geheilt wird. Die logit-Transformation ($\text{logit}(s) = \log(s/(1-s))$ für eine beliebige Zahl s) gewährleistet, dass die prognostizierten Werte des Modells alle innerhalb des Intervalls [0,1] liegen und als echte Wahrscheinlichkeiten interpretiert werden können. Für den Parameter β ergibt sich nach Exponentiation (also als $\exp(\beta)$) exakt das Odds Ratio aus der obigen Vierfeldertafel. Wie im linearen Regressionsmodell kann dieses Odds Ratio auch hier wieder als Steigungsparameter interpretiert werden. Bei Erhöhung der Kovariablen um eine Einheit, hier im nichtmetri-

schen Fall der Schritt von Kontrolle zu Verum, erhöht sich die Chance für das Eintreten des Zielereignisses (Heilung) um 50%.

Selbstverständlich kann das hier dargestellte logistische Regressionsmodell auch zu einem multiplen logistischen Regressionsmodell mit mehreren Kovariablen erweitert werden. Dabei bleiben alle Eigenschaften der multiplen Regressionsmodelle erhalten: Die Parameter können nach wie vor (nach Exponentiation) als Odds Ratios interpretiert werden, und der Einfluss einer jeden Kovariablen auf die Zielgröße ist jetzt um den Einfluss aller anderen Kovariablen im Modell bereinigt.

2. Regressionsmodelle für nichtstetige Zielgrößen und korrelierte Beobachtungen

In diesem Kapitel soll anhand des oben eingeführten Beispiels zum Vergleich zweier topischer Cremes das einfachste Regressionsmodell für nichtstetige Merkmale und korrelierte Beobachtungen vorgestellt werden, nämlich das für eine binäre Zielgröße. Es war bei der Vorstellung der Studie bereits angeklungen, dass diese eigentlich als multizentrische Studie durchgeführt worden war. Folgende Tabelle gibt einen Überblick über die Ergebnisse in den acht Studienzentren, wieder in der Form von einzelnen Vierfeldertafeln.

Zentrum		Heilung			OR [95%-KI]	Rel. Häufigkeit Heilung
		Nein	Ja			
1	Kontrolle	27	10	37	1,19 [0,43; 3,28]	29%
	Verum	25	11	36		
2	Kontrolle	10	22	32	1,82 [0,48; 6,85]	73%
	Verum	4	16	20		
3	Kontrolle	12	7	19	4,80 [1,20; 19,13]	55%
	Verum	5	14	19		
4	Kontrolle	16	1	17	2,29 [0,19; 27,99]	9%
	Verum	14	2	16		
5	Kontrolle	12	0	12	--	21%
	Verum	11	6	17		
6	Kontrolle	10	0	10	--	5%
	Verum	10	1	11		
7	Kontrolle	8	1	9	2,00 [0,10; 41,00]	14%
	Verum	4	1	5		
8	Kontrolle	1	6	7	0,33 [0,02; 5,03]	77%
	Verum	2	4	6		

Es ist auffällig, dass sich die Behandlungseffekte in den einzelnen Zentren sehr unterscheiden. Wir finden Zentren, in denen die Verumbehandlung der Kontrolle signifikant überlegen ist (Zentrum 3), daneben aber auch ein Zentrum (Zentrum 8), in dem die Kontrolltherapie besser als Verum abschneidet, wenn auch nicht statistisch signifikant. In zwei Zentren (5,6) kann das Odds Ratio aufgrund des Vorliegens einer Nullzelle gar nicht berechnet werden. Des Weiteren ist auffällig, dass die relative Heilungshäufigkeit (unabhängig vom Behandlungseffekt) sehr stark zwischen den Zentren schwankt, von nur 5% in Zentrum 6 bis zu 77% im Zentrum 8. Dies weist zum einen auf eine große Heterogenität der Patientenpopulationen in den einzelnen Zentren hin (offensichtlich findet man in Zentren mit geringerer Heilungshäufigkeit die kränkeren Patienten), andererseits auf mögliche Qualitätsunterschiede zwischen den Zentren. Unabhängig von den tatsächlichen Gründen erzeugt diese Heterogenität aber auch eine Korrelation zwischen den Patienten innerhalb der Zentren: In einem Zentrum mit hoher Heilungshäufigkeit werden viele Patienten geheilt und die Heilung eines einzelnen Patienten wird mit größerer Wahrscheinlichkeit von der Heilung eines anderen Patienten im selben Zentrum begleitet.

Die im vorigen Kapitel durchgeführte simple Analyse in einer einzelnen Vierfeldertafel berücksichtigt diese offensichtliche Korrelation jedoch nicht und wir werden im folgenden sehen, welche statistischen Methoden angewandt werden können, um diese Korrelation innerhalb der Beobachtungen zu berücksichtigen. Es wird dafür die folgende mathematische Notation verwendet: Es bezeichne Y_{ij} , ($i = 1, \dots, n$, $j = 1, \dots, n_i$), ob Patient j in Zentrum i geheilt wurde ($Y_{ij} = 1$: ja, $Y_{ij} = 0$: nein) und x_{ij} ob Patient j in Zentrum i in der Verum- oder in der Kontrollgruppe war ($x_{ij} = 1$: Verum, $x_{ij} = 0$: Kontrolle). Die Zielgröße Y_{ij} wird als Bernoulli-verteilt mit Erfolgswahrscheinlichkeit π_{ij} angenommen, das heißt $\pi_{ij} = P(Y_{ij} = 1) \sim B(1, \pi_{ij})$, der Erwartungswert von Y_{ij} ist dann ebenfalls gleich π_{ij} , $E(Y_{ij}) = \pi_{ij}$.

2.1 Marginale versus konditionale Modelle

Es gibt zwei große Klassen von Regressionsmodellen, die die Korrelation in den Beobachtungen adäquat berücksichtigen können, marginale und konditionale Modelle, wobei unter den konditionalen Modellen die Modelle mit zufälligen Effekten (auch: gemischte Modelle) die wichtigste Untergruppe darstellen^{18, 19}. Dieser Dichotomie

entsprechend haben die Parameter aus den beiden Modellklassen unterschiedliche Interpretationen.

2.1.1 Marginale Modelle

In einem marginalen Modell wird der Effekt der Kovariablen (hier: die Behandlung) getrennt von der Korrelation modelliert. Ein marginales logistisches Modell für unser Beispiel würde so aussehen:

$$\log(\pi_{ij} / (1 - \pi_{ij})) = \alpha + \beta x_{ij}$$

$$\text{Corr}(Y_{ij}, Y_{ij'}) = \gamma, \text{ für } j \neq j'$$

$$\text{Corr}(Y_{ij}, Y_{i'j'}) = 0, \text{ für } i \neq i'$$

Die erste Gleichung unterscheidet sich von der Gleichung des herkömmlichen Modells nur dadurch, dass jetzt die beiden Ebenen im Datensatz (Patienten und Zentren) explizit unterschieden werden müssen. Die zweite Gleichung besagt, dass die Heilung zweier Patienten j und j' innerhalb eines Zentrums korreliert ist und diese Korrelation gleich γ ist. Diese Korrelation wird als konstant zwischen allen Patienten innerhalb eines Zentrums und als konstant über alle Zentren angenommen. Die dritte Gleichung schließlich bildet die Annahme ab, dass die Heilungen von Patienten aus unterschiedlichen Zentren unkorreliert sind.

Die Interpretation der uns eigentlich interessierenden Kovariablen ist analog zum herkömmlichen logistischen Regressionsmodell. Der transformierte Regressionskoeffizient $\exp(\beta)$ ist gleich der Chance auf Heilung in der Verumgruppe geteilt durch die Chance auf Heilung in der Kontrollgruppe. Die Interpretation dieses Koeffizienten ist unabhängig vom jeweiligen Zentrum bzw. gilt für die ganze Population von Zentren. Genauer gesagt, der Behandlungseffekt wird über die ganze Gruppe von Zentren gemittelt, daher werden die Koeffizienten von marginalen Modellen auch häufig als „population-averaged“ bezeichnet. Wird in allen Zentren mit Verum anstatt mit Kontrolle behandelt, dann steigt die mittlere Chance auf Heilung um $\exp(\beta)$. Die Parameterschätzung für marginale logistische Regressionsmodelle erfolgt meistens mit Hilfe der so genannten GEE (Generalized estimation equations) Methode, die ursprünglich von Liang/Zeger²⁰ vorgeschlagen wurde. Für unser Beispiel ergibt sich ein geschätztes Odds Ratio ($\exp(\beta)$) von 1,74 [1,10; 2,75], der zugehörige p-Wert ist $p = 0,017$. Es fällt auf, dass der Behandlungseffekt jetzt (nach Berücksichtigung der Korrelation in-

nerhalb der Zentren) statistisch signifikant erhöht ist, auch der geschätzte Wert des Odds Ratios hat sich von 1,50 auf 1,74 erhöht.

2.1.2 Konditionale Modelle

In einem konditionalen Modell (wobei wir uns hier auf die Modelle mit zufälligen Effekten beschränken) wird angenommen, dass es eine natürliche Heterogenität zwischen den Zentren gibt und dass diese Heterogenität mit Hilfe einer Wahrscheinlichkeitsverteilung modelliert werden kann. Das heißt in unserem Fall, dass die zugrunde liegenden Heilungswahrscheinlichkeiten von Zentrum zu Zentrum verschieden sind, diese aber insgesamt einer Normalverteilung folgen. Die Korrelation der Patienten innerhalb eines Zentrums entsteht also dadurch, dass den Patienten spezifische, aber nicht beobachtete (Umgebung, Personal etc.) Eigenschaften des Zentrums gemeinsam sind. Der eigentlich interessierende Behandlungseffekt wird aber weiterhin als konstant über alle Kliniken angenommen. Als Modellgleichung für solch ein logistisches Regressionsmodell mit zufälligem Intercept (konstanten Faktor) ergibt sich

$$\log(\pi_{ij} / (1 - \pi_{ij}) | u_i) = \alpha + \beta x_{ij} + u_i$$

mit $u_i \sim N(0, \sigma^2)$. Es wird weiter angenommen, dass, gegeben u_i , die Ereignisse innerhalb eines Zentrums unabhängig sind. Das heißt, dass die Korrelation zwischen zwei Patienten in einem Zentrum vollständig dadurch erklärt wird, dass diese im selben Zentrum behandelt worden sind. Dadurch dass nur der konstante Faktor als zufällig angenommen wird, der Behandlungseffekt aber nach wie vor als fester Effekt, wird von einem über alle Zentren konstanten Behandlungseffekt ausgegangen. Über die Zentren variiert also nicht der Behandlungseffekt, sondern nur die zugrunde liegende Heilungswahrscheinlichkeit. Aufgrund dieser zentren-spezifischen Heilungswahrscheinlichkeiten werden die Parameter aus diesen Modellen auch häufig als „subject-specific“ bezeichnet, wobei mit „subjects“ hier die Zentren gemeint sind. Die Mischung aus zufälligen und festen Effekten im Modell erklärt darüber hinaus auch die häufig benutzte Bezeichnung „gemischte Modelle“.

Auch im konditionalen Modell ist die Interpretation der Parameter wieder analog zu den vorherigen logistischen Regressionsmodellen durchzuführen. Der transformierte Regressionskoeffizient $\exp(\beta)$ ist gleich der Chance auf Heilung in der Verumgruppe

geteilt durch die Chance auf Heilung in der Kontrollgruppe. Es gibt jedoch subtile Unterschiede. Während in marginalen Modellen der Behandlungseffekt marginal zu interpretieren ist (wie ändert sich die Heilungswahrscheinlichkeit durch die Behandlung, gemittelt über alle Zentren), ist der Behandlungseffekt in konditionalen Modellen bedingt auf das jeweilige Zentrum zu interpretieren: Wie ändert sich die Heilungswahrscheinlichkeit durch die Behandlung in einem Zentrum mit dem Wert u_i ? Die Varianz σ^2 misst den Grad der Heterogenität zwischen den Zentren, je größer σ^2 , desto größer sind die Unterschiede zwischen den Zentren. Entsprechend der obigen Überlegung wird über σ^2 auch die Stärke der Korrelation innerhalb der Zentren gesteuert. Je größer die Heterogenität zwischen den Zentren, desto größer auch die Korrelation innerhalb der Zentren. Die konkrete Berechnung dieser so genannten „Intra-Class-Korrelation“ (ICC) ist im binären Fall allerdings nicht trivial²¹.

Es existiert eine Fülle von Methoden für die Parameterschätzung in logistischen Regressionsmodellen mit zufälligen Effekten. Diese Vielzahl der Möglichkeiten ist auf die Komplexität der Schätzung zurückzuführen. Das Problem dabei ist, die die Likelihood-Funktion aus einem Produkt von n (= der Anzahl von Zentren) Integralen besteht, von denen jedes einzelne jedoch nicht geschlossen lösbar ist. Daher wurden eine Fülle von Approximationsmethoden^{22, 23} entwickelt, eine exakte Auswertung der Likelihood-Funktion ist nur mit Methoden der numerischen^{18, 24, 25} oder der stochastischen Integration²⁶ (MCMC) möglich.

In unserem Beispiel ergeben sich als geschätzte Werte für die Parameter 2,09 [1,16; 3,77], $p = 0,014$ für das Odds Ratio der Behandlung und 1.96 [0; 4,77] für σ^2 , wobei eine numerische Integrationsmethode (Adaptive Gauß-Quadratur, SAS PROC NLMIXED) angewandt wurde und die untere Grenze des Konfidenzintervalls für σ^2 auf 0 gesetzt wurde, da negative Werte für eine Varianz nicht definiert sind. Es fällt auf, dass der geschätzte Behandlungseffekt im gemischten Modell größer ist als der im marginalen Modell. Es kann gezeigt werden, dass dies notwendigerweise der Fall ist²⁷, im vorliegenden Fall kann sogar ein approximativer Umrechnungsfaktor angegeben werden²⁸.

3. Anwendungen von Regressionsmodellen für nichtstetige Zielgrößen und korrelierte Beobachtungen in der Medizin

Regressionsmodelle erfreuen sich immer größerer Beliebtheit und Anwendung im medizinischen Bereich. Vittinghoff et al.²⁹ behaupten im Vorwort ihres Lehrbuches zu Regressionsmodellen in der Biostatistik sogar, dass Regressionsmodelle das wichtigste statistische Werkzeug in der modernen medizinischen Forschung sind. Nun ist ein solcher Satz aus Werbezwecken im Vorwort eines Lehrbuches durchaus zu erwarten, aber die immer häufigere Anwendung von Regressionsmodellen ist auch empirisch zu belegen.

Bei einer Analyse der verwendeten statistischen Methoden in den Jahrgängen 2004 und 2005 des „New England Journal of Medicine“³⁰ findet sich in 51% der Veröffentlichungen ein multiples Regressionsmodell, in weiteren 7% ein einfaches lineares Regressionsmodell.

Interessanterweise findet sich in diesem Zeitraum auch in 12% der Veröffentlichungen eine statistische Analyse, die explizit das Auftreten von korrelierten Beobachtungen berücksichtigt. In den Jahren 1978-79 und 1989 war solch eine Analyse in keinem einzigen Fall aufgetaucht.

Die Entwicklung zu immer komplexeren statistischen Methoden ist nicht nur auf das NEJM beschränkt. Auch in einer Analyse der Zeitschrift „Arthritis & Rheumatism“ (Impact-Factor 2006: 7,7) finden sich in 19% der Veröffentlichungen im Jahre 2005 Regressionsmodelle und in 10% der Fälle Modelle für korrelierte Daten (GEE, Gemischte Modelle und Repeated-Measurement ANOVA)³¹.

Dieser zunehmenden Komplexität entspricht nicht immer das entsprechende Wissen um die Zusammenhänge bei den Lesern. In einer Befragung von amerikanischen Assistenzärzten und -ärztinnen in der Inneren Medizin³² waren nur 13% in der Lage, in einer MC-Frage korrekt ein Cox-Modell zu identifizieren. Nur 37% war es gelungen, ein Odds Ratio aus einem logistischen Regressionsmodell korrekt zu interpretieren.

Ein Indiz für die zunehmende Relevanz von Regressionsmodellen in der medizinischen Forschung ist auch die Publikation von zahlreichen Tutorials und nichttechnischen Übersichtsartikeln im medizinischen Schrifttum. In der Reihe „Fundamentals of Clinical Research for Radiologists“ des American Journal of Roentgenology findet sich ein Artikel zur Regression³³ und einer zu Modellen mit multivariaten Zielgrö-

ßen³⁴. Die Zeitschrift „Circulation“ hat in ihrem „Statistical Primer for Cardiovascular Research“ einen Artikel über Regressionsmodelle eingeschlossen³⁵. „Clinical Orthopaedics and Related Research“ nennt in einem Übersichtsartikel³⁶ über statistische Methoden Regressionsmodelle „important statistical concepts for clinicians involved in the conduct or appraisal of orthopaedic clinical research“. In einem nur 10-seitigen Artikel werden den Regressionsmodellen zwei Seiten gewidmet und es werden lineare, logistische und Cox-Modelle beschrieben. Die „Annals of Thoracic Surgery“ widmen den Regressionsmodellen auf ihrer „Statistician’s page“ zwei Artikel^{37, 38} mit besonderem Schwerpunkt auf den logistischen Modellen. Das „British Journal of Cancer“ beleuchtet in einer Serie zur „Survival Analysis“ das Cox-Modell in zwei Artikeln^{39, 40}. Auch in den „Annals of Internal Medicine“ findet sich ein Tutorial zur multiplen Regressionsanalyse⁴¹. Die Zeitschrift „Radiology“ unterscheidet in ihrer Serie zu „Statistical Concepts“ explizit zwischen der simplen⁴² und der multiplen Regression⁴³ und widmet beiden Themen jeweils einen Artikel.

Auch die deutschsprachigen medizinischen Zeitschriften haben erkannt, wie wichtig klare und verständliche Einführungen in statistische Regressionsmodelle sind. Die „Deutsche Medizinische Wochenschrift“ widmet in ihrer 29-teiligen „Statistik“-Serie den verschiedenen Regressionsmodellen gleich vier Beiträge⁴⁴⁻⁴⁷.

Im Bereich der korrelierten Beobachtungen findet sich ebenfalls eine Reihe von Übersichtsartikeln über statistischen Methoden für Nichtstatistiker. In der wahrscheinlich bekanntesten und umfangreichsten Serie zu statistischen Methoden, der des „British Medical Journal“, wird in mindestens sechs Artikeln auf die Probleme von korrelierten Daten eingegangen.⁴⁸⁻⁵³ Das „Journal of Thoracic and Cardiovascular Surgery“ bezeichnet in einem Artikel seiner Serie „Statistics for the Rest of Us“ die Möglichkeit der Analyse von wiederholten Messungen als einen Durchbruch auf einem Weg zu effektiven, hilfreichen und korrekten Analysen von klinischen Outcomes⁵⁴. Auch im Bereich der korrelierten Beobachtungen gibt es bereits deutschsprachige Übersichtsartikel, eines in der bereits erwähnten „Statistik“-Serie der DMW⁵⁵ und eines in der „Methoden“-Reihe der Zeitschrift „Rehabilitation“⁷.

Im folgenden werden vier konkrete Anwendungen von Regressionsmodellen für nichtstetige Zielgrößen und korrelierte Beobachtungen dargestellt^{4, 56-58}. Die Volltexte der Publikationen zu den Auswertungen der besprochenen Studien finden sich im Anhang. Es werden dabei binäre, nominale und diskrete Zielgrößen betrachtet, außerdem variiert das Anwendungsgebiet von der Kardiochirurgie über die Onkologie

und Versorgungsforschung zur Allergologie. Auch der jeweilige methodische Fokus der Anwendungsbeispiele ist unterschiedlich, es werden Fragen der Schätzmethoden, der zu verwendenden Software und zu Clusterungseigenschaften der Modelle besprochen.

3.1 Ein exakter Test für die Meta-Analyse mit binären Endpunkten

Wir haben im Kapitel 2 gesehen, dass die Parameterschätzung in Regressionsmodellen für nichtstetige Zielgrößen und korrelierte Beobachtungen anspruchsvoll ist, sowohl was die zugrundeliegende statistische Theorie als auch die Software angeht. Bereits ein logistisches Regressionsmodell mit einem normalverteilten zufälligen Intercept verlangt ein numerisches Integrationsverfahren, wenn man auf echten (im Sinne von nichtapproximativen) Maximum-Likelihood-Schätzern besteht. Schon die Parameterschätzung für ein herkömmliches logistisches Regressionsmodell (mit unkorrelierten Beobachtungen) ist nur mit Hilfe eines iterativen Algorithmus zu bewerkstelligen.

Es überrascht daher, dass es einen Spezialfall eines Regressionsmodells für nichtstetige Zielgrößen und korrelierte Beobachtungen gibt, in dem nicht nur eine geschlossene Schätzung der Parameter möglich ist, sondern sogar exakte (im Sinne von nichtasymptotische) Schätzer berechnet werden können. Bei diesem Spezialfall handelt es sich um die Meta-Analyse mit binären Endpunkten, d.h. das Ziel ist die numerische Zusammenfassung des Behandlungseffekts aus K verschiedenen (in der Regel randomisierten) Studien, in denen jeweils eine Verumbehandlung mit einer Kontrollbehandlung verglichen wird und die Zielgröße binär ist. Dieses Design kann in der Sprache der log-linearen Modelle als $2 \times 2 \times K$ -Tafel betrachtet werden und es ist im Prinzip seit langem bekannt^{59, 60}, wie statistische Tests und Parameterschätzer in diesem Fall exakt berechnet werden können. Das zugrundeliegende Prinzip bei den exakten Verfahren ist dabei wieder das der Konditionierung, und zwar auf eine suffiziente Statistik. In der oben eingeführten Unterscheidung zwischen marginalen und konditionalen Modellen handelt es hierbei also um ein konditionales Modell. Allerdings werden hier keine zufälligen Effekte in die Modellgleichung aufgenommen, stattdessen wird der Studieneffekt komplett aus der Likelihood-Funktion des Modells „herausbedingt“, d.h. es wird eine konditionale Likelihood-Funktion maximiert. Eine Approximation an diese konditionale Likelihood-Funktion führt zum herkömmlichen

(asymptotischen) Mantel-Haenszel-Test⁶¹, der sich in der Meta-Analyse bereits bewährt hat. Da exakten Verfahren immer noch große Anforderungen an die Rechenzeit stellen, werden solche Verfahren vor allem in der Meta-Analyse von seltenen Ereignissen oder unerwünschten Nebenwirkungen von Interesse sein.

Es ist überraschend, dass diese Erkenntnis im Bereich der Meta-Analyse nicht sehr verbreitet ist, in den beiden Standardlehrbüchern zur Meta-Analyse^{62, 63} findet sich nur eine einzelne Erwähnung von exakten Verfahren und dabei der Verweis auf das Paper von Emerson⁵⁹.

Es sei darauf hingewiesen, dass auch die Studie von Beitler/Landis aus der Einleitung als solches Problem interpretiert werden kann, hierfür ergibt sich ein geschätztes exaktes Odds Ratio von 2,13 [1,14; 4,08], der zugehörige p-Wert ist $p = 0,016$.

Inwieweit das „wiederentdeckte“ exakte Verfahren tatsächlich eine Verbesserung gegenüber den Standardmethoden der Meta-Analyse (Modell mit festen oder mit zufälligen Effekten) und dem asymptotischen Mantel-Haenszel-Test darstellt, wurde in einer Simulationsuntersuchung geprüft, mit besonderem Fokus auf Situationen mit seltenen Ereignissen⁵⁶. Dabei zeigte sich, dass in nahezu allen untersuchten Situationen der asymptotische Mantel-Haenszel-Test den anderen Verfahren überlegen ist, vor allem die Standardmethoden der Meta-Analyse schneiden hier schlecht ab. Insofern besteht keine Notwendigkeit das exakte Verfahren einzusetzen, der asymptotische Mantel-Haenszel-Test reicht in der Regel zur Beurteilung des Behandlungseffektes in einer Meta-Analyse mit binären Endpunkten vollkommen aus. Die Standardmethoden (Modell mit festen oder mit zufälligen Effekten) sollten bei seltenen Ereignissen dagegen nicht eingesetzt werden.

Dass diese Erkenntnisse für die klinische Anwendung tatsächlich relevant sind, zeigt eine Meta-Analyse zum Vergleich der Off- und On-Pump-Technik (ohne und mit Herz-Lungen-Maschine) in der Bypasschirurgie⁶⁴. Hier waren in 21 Studien insgesamt nur 20 postoperative Schlaganfälle beobachtet worden.

Methode	Odds Ratio [95%-KI] Off-Pump vs. On-Pump	p-Wert
Modell mit festen Effekten	0,54 [0,22; 1,33]	0,168
Modell mit zufälligen Effekten	0,67 [0,33; 1,39]	0,286
Mantel-Haenszel-Verfahren	0,43 [0,17; 1,14]	0,081
Exaktes Verfahren	0,44 [0,14; 1,21]	0,094

Es zeigt sich, dass die Standardmethoden den protektiven Effekt der Off-Pump-Technik unterschätzen, sowohl was den geschätzten Effekt, als auch den zugehörigen p-Wert angeht. Eine eigene spätere Analyse mit einem größeren Datensatz⁶⁵ zeigt tatsächlich einen signifikanten Effekt (OR: 0,38 [0,18; 0,81], $p = 0,009$).

3.2 Ein Regressionsmodell für die relative Überlebenszeit bei korrelierten Beobachtungen

Zur Analyse des tumorspezifischen Überlebens in populationsbasierten Tumorregistern wird inzwischen fast routinemäßig die relative Überlebenszeitanalyse herangezogen⁶⁶. Das Prinzip dabei ist, das Überleben der Registerpatienten am Überleben der Allgemeinbevölkerung (basierend auf den amtlichen Sterbestatistiken) zu standardisieren⁶⁷. Dadurch erhält man ein Maß für die Übersterblichkeit in der Gruppe der Registerpatienten, unabhängig davon, ob der Tumor die tatsächliche Todesursache ist oder ob dieser auf dem Totenschein als Todesursache angegeben ist⁶⁸. Informationen über die Todesursachen der Register-Patienten, die oft mit Mängeln behaftet sind, werden so nicht mehr benötigt. Des Weiteren kann das relative Überleben als das Überleben in einer hypothetischen Population interpretiert werden, in der der interessierende Tumor die einzige Todesursache ist und man erhält so ein Maß für das Quasi-todesursachenspezifische Überleben⁶⁹. Die Qualität dieser Maße ist allerdings davon abhängig, wie gut das erwartete Überleben geschätzt worden ist und ob die Patientengruppe eine Stichprobe aus der Allgemeinbevölkerung ist. Die relative Überlebenszeitanalyse hat noch weitere Vorteile. Zum einen erlaubt sie Aussagen über die Heilung (im statistischen, nicht im klinischen Sinne) von der Tumorerkrankung, nämlich dann wenn das beobachtete Überleben in der Gruppe der Registerpatienten gleich dem erwarteten Überleben in der Allgemeinbevölkerung ist. Zum

anderen werden internationale Vergleiche zwischen Tumorregistern erleichtert, wenn die Registerpatienten aus verschiedenen Ländern an der jeweiligen Allgemeinbevölkerung standardisiert worden sind.

Über die reine Beschreibung des relativen Überlebens hinausgehend, wurde eine Reihe von Regressionsmodellen für das relative Überleben vorgeschlagen^{68, 70, 71}, um den Einfluss von prognostischen oder Risikofaktoren auf das relative Überleben zu messen. Aus der Definition des relativen Überlebens als einem Quotienten folgt unmittelbar, dass es sich hierbei um Modelle handeln muss, die additiv in den Hazards sind. Zum Vergleich, das Cox-Modell für herkömmliche Überlebenszeiten ist multiplikativ in den Hazards. Das Modell von Estève⁶⁸, das als einziges der genannten Modelle die exakten beobachteten Überlebenszeiten berücksichtigt, kann dabei als Generalisiertes Lineares Modell¹⁰ mit einer binären Zielgröße, einer Poisson-Likelihood, einem Offset und einer spezifischen Link-Funktion, die für jeden Patienten unterschiedlich ist, interpretiert werden⁷². Eine Verallgemeinerung eines Regressionsmodells für das relative Überleben für korrelierte Beobachtungen kann dann leicht angegeben werden, in dem dieses Generalisierte Lineare Modell in die Familie der Generalisierten Linearen Gemischten Modelle (GLMM) eingebettet wird⁴. Dabei wird, ganz ähnlich wie im oben beschriebenen logistischen Regressionsmodell mit einem zufälligen Intercept, ein solch zufälliger Intercept in die Modellgleichung eingesetzt, um die Korrelation mitzumodellieren.

Eingesetzt haben wir dieses erweiterte Modell bei der Analyse des relativen Überlebens in der HALLUCA-Studie⁷³, einer versorgungsepidemiologischen Studie zum Lungenkrebs in der Region Halle. In enger Zusammenarbeit mit den regionalen Tumorzentren waren dazu von April 1996 bis September 1999 alle an Lungenkrebs Erkrankten in der Region Halle registriert worden. Die Nachbeobachtung der Patienten erfolgte bis September 2000. Insgesamt waren 1696 Lungenkrebspatienten beobachtet worden, 1349 (79,5%) waren bis zum Studienende verstorben, die mediane Überlebenszeit hatte 284 Tage (9,3 Monate) betragen. Bereits bei der rein deskriptiven Analyse der Studie hatten wir sehr heterogene Überlebenszeiten in den 56 verschiedenen diagnostischen Zentren der Studienregion beobachtet. Diese Unterschiede sind aus klinischer Sicht durchaus plausibel. Es ist zu erwarten, dass niedergelassenen Ärzte, die die Verdachtsdiagnose „Lungenkrebs“ stellen, ihre Patienten dann zur diagnostischen Abklärung an diejenigen Einrichtungen überweisen, die auch adäquate Behandlungsmöglichkeiten haben.

Um diese Unterschiede zwischen den Zentren bei der Analyse des relativen Überlebens mit zu berücksichtigen, haben wir die oben beschriebene Erweiterung des Estève-Modells für korrelierte Beobachtungen in ein GLMM mit Poisson-Likelihood, einem Offset, einer spezifischen Link-Funktion und einem zufälligen Intercept angewandt. Es zeigte sich, dass die Parameterschätzer für die Risikofaktoren im Vergleich zum Standard-Estève-Modell für unkorrelierte Beobachtungen nur unwesentlich anders waren. Jedoch hatte ein Vergleich der BIC-Werte zwischen den Modellen einen klaren Vorteil für das erweiterte Modell erbracht, auch die Varianz des zufälligen Intercepts war statistisch signifikant gewesen ($p = 0,028$).

3.3 Die Schätzung von nominalen logistischen Modellen mit korrelierten Beobachtungen in SAS

Wie wir bisher gesehen haben, sind Regressionsmodelle für korrelierte Beobachtungen mit binären und diskreten Zielgrößen noch relativ beherrschbar, vor allem deshalb, weil sich diese in die Klassen der Generalisierten Linearen Gemischten Modelle (GLMM) einbetten lassen, für die sowohl die mathematische Theorie als auch die Entwicklung der Software für die Parameterschätzung weit fortgeschritten ist. In der im professionellen Bereich (z.B. Pharmaindustrie) am weitesten verbreiteten Software SAS steht z.B. dafür eine eigene Prozedur (PROC GLIMMIX) zur Verfügung, auch mit anderen Prozeduren (PROC NLMIXED, PROC GENMOD) können Regressionsmodelle für binäre und diskrete Zielgrößen mit korrelierten Beobachtungen relativ leicht geschätzt werden. In einer schon nahezu unübersichtlichen Vielfalt stehen zudem aus dem mathematisch-theoretischen Bereich zusätzliche Schätzverfahren zur Verfügung, die aber nur selten in allgemein zugänglicher Software umgesetzt sind. Beschränkt man sich nur auf die Methoden für binäre Zielgrößen, die auch mit einem kurzen Akronym eingeführt oder genutzt worden sind, findet man GEE⁷⁴, CGEE⁷⁵, EGEE⁷⁶, REGEE⁷⁷, OCEE⁷⁸, GQL⁷⁹, GGEE⁷⁹, GEGEE⁷⁹, EEE⁸⁰ und QIF⁸¹ als Erweiterungen oder Varianten der ursprünglichen GEE-Methode. Für konditionale Modelle ist die Liste noch länger. MQL²², PQL²², MQL⁸², MCEM⁸³, MCNR⁸³, SML⁸³, MSM⁸⁴, HGLM⁸⁵, DHGLM⁸⁶, SA-MCMC⁸⁷, SA-EM⁸⁸, QMC⁸⁹, GHQ⁹⁰, NPMLE⁹¹, ROBMS-GLMM⁹², MPLE⁹³, MMM⁹⁴, GNMM⁹⁵, RIGLS⁹⁶, AGQ⁹⁷, Laplace⁹⁸, RMNCR⁹⁹, Slice-EM1¹⁰⁰, Slice-EM2¹⁰⁰, CGLS¹⁰¹, QELS¹⁰² und REMQL¹⁰³ wurden vorgeschlagen, um die Parameter eines logistischen Regressionsmodells mit einem zufälligen Intercept zu schätzen. Nur ein verschwindend klei-

ner Teil dieser Verfahren ist in den Standardprozeduren von SAS umgesetzt, SPSS als ebenfalls häufig verwendete Software im medizinischen Bereich bietet *kein einziges* dieser Verfahren an, auch nicht GEE oder PQL. Auch wenn eine Reihe von Autoren Makros oder R-Prozeduren zur Verfügung stellt, ist offensichtlich, dass eine riesige Diskrepanz zwischen den vorgeschlagenen Verfahren und deren Umsetzbarkeit in den Standard-Software-Paketen besteht.

Schwieriger wird die Situation, wenn man sich für Modelle mit echt nominalen Zielgrößen und korrelierten Beobachtungen interessiert, wo also die Zielgröße eine endliche Anzahl von Kategorien R besitzt, die aber echt größer als 2 ist. In diesen Modellen muss eine der R Kategorien der Zielgröße als Referenzkategorie ausgewählt werden, für jede Kovariable werden dann $R-1$ Parameterschätzer ausgegeben, die alle jeweils im Vergleich mit der Referenzkategorie interpretiert werden müssen. Dementsprechend müssen auch für den einfachsten Fall eines zufälligen Intercepts bereits $R-1$ Modellgleichungen angegeben werden:

$$\log(\pi_{ijr} / (\pi_{ij1}) | u_{ir}) = \alpha_r + \beta_r x_{ij} + u_{ir}, \quad r = 2, \dots, R.$$

Für die zufälligen Intercepts u_{ir} muss dabei eine multivariate Normalverteilung mit Erwartungswert 0 und Kovarianzmatrix Σ angenommen werden. Hartzel et al.¹⁰⁴ haben gezeigt, wie ein solches Modell durch adaptive Gauss-Hermite-Quadratur mit PROC NLMIXED in SAS geschätzt werden kann.

Einen anderen Zugang haben Chen/Kuo¹⁰⁵ und in deren Nachfolge Malchow-Møller/Svarer¹⁰⁶ gewählt. Beide haben den wohlbekannten Zusammenhang zwischen einer Multinomial- und einer Poisson-Verteilung genutzt (siehe z.B. McCullagh/Nelder¹⁰⁷), um ein nominales Regressionsmodell mit zufälligen Intercepts als Poisson-Modell zu schreiben und zu schätzen. Dieses Vorgehen ist mathematisch vollkommen korrekt, führt aber zu einer beträchtlichen Erhöhung der Rechenzeit und zu einer Instabilität der Schätzalgorithmen, da eine größere Anzahl von Parametern geschätzt werden muss.

Wir haben auf diese Tatsachen hingewiesen und gleichzeitig gezeigt, wie Regressionsmodelle mit nominalen Zielgrößen und korrelierten Beobachtungen in SAS geschätzt werden können⁵⁷. Dabei haben wir die Idee von Hartzel et al.¹⁰⁴ noch einmal aufgenommen, aber zusätzlich auch eine PQL-Schätzung mit PROC GLIMMIX und eine GEE-Schätzung für ein marginales Modell mit PROC GENMOD vorgestellt. In den beiden letzten Fällen ist eine Schätzung mit SAS nur möglich, wenn man das

Modell mit Hilfe des „Wright“-Tricks¹⁰⁸ als ein echtes multiples Modell schreibt. Das heißt, die gegenwärtigen (Version 9.1.3) Implementationen von PROC GENMOD und PROC GLIMMIX schätzen kein echtes nominales Modell für korrelierte Beobachtungen, obwohl beide Prozeduren scheinbar solche Optionen anbieten.

Die Motivation zur Beschäftigung mit diesem Problem hatte eine Studie geliefert, die sich mit Faktoren beschäftigte, die die Einschätzung der optimalen rehabilitativ-medizinischen Versorgung von Patienten mit Schädel-Hirn-Trauma (SHT) beeinflussen¹⁰⁹. Dazu waren in einer Studie des Instituts für Neurologisch-Neurochirurgische Rehabilitationsforschung (INNRF) Magdeburg einer Reihe von Ärzten/Ärztinnen 10 fiktive, aber realitätsnahe Musterkrankengeschichten (MKG) von Schädel-Hirn-Trauma-Patienten vorgelegt worden. Die Ärzte hatten dann ein optimales Rehabilitationssetting (stationär, teilstationär oder ambulant) für jeden dieser Musterfälle festzulegen. Es sollte nun geprüft werden, ob einerseits Eigenschaften der beurteilenden Ärzte/Ärztinnen (Fachrichtung, Ausbildungsstand), andererseits Unterschiede der beschriebenen Musterkranken (Verlauf der Krankheit, Grad der Beeinträchtigung) die Einschätzung beeinflussen. Dabei ist offensichtlich, dass die Einschätzungen innerhalb der Ärzte/Ärztinnen als korreliert angenommen werden müssen. Im Ergebnis zeigte sich, dass (wie aus Patientensicht wohl erhofft) die Arztmerkmale die Setting-Präferenzen kaum beeinflussen. Die beiden Kovariablen, die die jeweilige MKG beschreiben, waren dagegen hochsignifikant. Je schwerer die Beeinträchtigung und je kürzer das SHT zurückliegt, um so eher wird eine stationäre Rehabilitation präferiert. Bei genauerer Betrachtung der Zielgröße „Optimales Rehabilitationssetting“ fällt auf, dass diese durchaus auch als ordinale Zielgröße interpretiert werden könnte, z.B. als grobe Kategorisierung einer latent zugrundeliegenden stetigen Zielgröße „Verbrachte Zeit in der Reha-Einrichtung“. In einer anderen Publikation haben wir diesen Datensatz genutzt, um ein Stereotype Regressionsmodell^{110, 111} für korrelierte Beobachtungen vorzuschlagen¹¹². Der Vorteil dieses Stereotype-Modells gegenüber dem nominalen Modell ist der, dass Aussagen über eine mögliche Anordnung der Kategorien der Zielgröße möglich sind und dafür auch eine Abstandsmetrik definiert ist. Als Ergebnis für den SHT-Datensatz zeigte sich, dass sich das teilstationäre Setting ungefähr in der Mitte zwischen dem stationären und dem ambulanten Setting einordnet.

3.4 Datenabhängiges, modellbasiertes Gruppieren von binären longitudinalen Verläufen am Beispiel der Neurodermitis

Wenn wir bisher konditionale Modelle oder Modelle mit zufälligen Effekten betrachtet haben, waren wir stets davon ausgegangen, dass die zufälligen Effekte als stetig und normalverteilt vorausgesetzt werden können. Diese Annahme ist durchaus plausibel, wenn man den zufälligen Effekt pro Cluster als eine Summe von vielen unbekanntem Merkmalen mit jeweils unabhängigen kleinen Effekten auffasst. Daran garantiert nämlich der zentrale Grenzwertsatz eine Normalverteilung dieser Summe.

Manchmal ist es jedoch plausibler anzunehmen, dass sich die Gesamtheit der Cluster aus L verschiedenen unterschiedlichen Gruppen von Clustern zusammensetzt. Die zufälligen Effekte über die Cluster sind dann besser durch eine diskrete Verteilung mit L Stützstellen, eine so genannte Mischverteilung, zu beschreiben. Die Parameterschätzung in solchen Modellen wird dann allerdings dadurch kompliziert, dass zusätzlich zu den festen Effekten noch die komplette Mischverteilung geschätzt werden muss, insbesondere auch die Anzahl L der Stützstellen. Zum Vergleich, im herkömmlichen Modell mit zufälligen Effekten muss zusätzlich zu den festen Effekten nur ein einzelner zusätzlicher Parameter, nämlich die Varianz des zufälligen Effektes, geschätzt werden.

Ein großer Vorteil der Mischungsmodelle ist der, dass für jedes Cluster auch Wahrscheinlichkeitsaussagen über die Zugehörigkeit zu den jeweiligen Gruppen möglich sind und sich dadurch auf natürliche Art und Weise Gruppen von Clustern bilden. Im einfachsten Fall einer binären Zielgröße hat Aitkin⁹¹ gezeigt, wie ein iterativer Algorithmus von gewichteten logistischen Regressionsmodellen benutzt werden kann, um ein solches Mischungsmodell zu schätzen. Für unseren Datensatz zum Vergleich zweier Cremes zur Behandlung nichtspezifischer Infektionen (wo die Zentren jetzt die einzelnen Cluster darstellen) ergibt sich ein minimaler Wert der Likelihood-Funktion (als Optimalitätskriterium für die Stützstellen) für ein Modell mit $L = 4$ Stützstellen. Der Schätzer (in gewohnter Weise als Odds Ratio) für den Behandlungseffekt ist dabei 2,02 [1,14; 3,58], der zugehörige p -Wert ist $p = 0,017$. Die Wahrscheinlichkeit der Gruppenzugehörigkeiten ist in unten stehender Tabelle dargestellt, die jeweils höchste Wahrscheinlichkeit pro Zentrum ist grau unterlegt. In Gruppe 1 finden sich also die Zentren 4, 5, 6 und 7, in Gruppe 2 nur das Zentrum 1 und in Gruppe 4 die Zentren 2,3, und 8. Die Gruppenzugehörigkeiten definieren sich also v.a. durch die unter-

schiedlichen Baseline-Heilungsraten. In Gruppe 1 finden sich die Zentren mit einer niedrigen, in Gruppe 4 die Zentren mit einer hohen Baselineheilungsrate, in der Gruppe 2 findet sich ein Zentrum mit mittlerer Baselineheilungsrate.

Zentrum	Wahrscheinlichkeit der Zugehörigkeit (in %) zu Gruppe ...			
	1	2	3	4
1	8,3	51,0	40,8	0,0
2	0,0	0,0	0,0	100,0
3	2,1	16,5	29,4	52,0
4	91,6	5,0	3,4	0,0
5	57,7	22,7	19,5	0,1
6	96,8	2,8	0,4	0,0
7	49,8	24,9	24,9	0,4
8	1,9	10,1	10,1	78,0

Eine sehr breite Anwendung solcher Mischungsmodelle findet sich in der Kriminologie bei der Beschreibung von kriminellen Karrieren von Heranwachsenden. Die Initialzündung war dabei die Publikation von Nagin¹¹³, wo ein solches Mischungsmodell mit einer gründlichen Beschreibung der Annahmen, aber auch der Vorteile des Modells vorgestellt worden ist.

Wir haben dieses Modell angewandt⁵⁸, um Neurodermitisverläufe von Kindern zu beschreiben. Die Neurodermitis ist eine Hautkrankheit, die bevorzugt im Kindesalter auftritt, eine starke genetische Komponente hat und sich chronisch rezidivierend typischerweise durch Hauteckzeme mit quälendem Juckreiz zeigt. Die Prävalenz der Neurodermitis bei Kindern ist in den Industrieländern hoch und steigt beständig an. Der Verlauf der Krankheit ist wechselhaft und verläuft schubweise, ist aber praktisch nicht zu prognostizieren, Zeiten mit in unregelmäßigen Abständen auftretenden Schüben wechseln mit Zeiten der Abheilung. Insbesondere ist völlig unklar, ob es charakteristische Gruppen von Verläufen gibt, die bereits bei der Geburt ausgehend von den vorliegenden endogenen und exogenen Risikofaktoren eine Prognose des Verlaufs erlauben würden. Für unsere Analyse lagen die Verläufe von 1314 Kindern aus der MAS-Studie¹¹⁴ vor. Dort war an 10 Messzeitpunkten zwischen dem 3. Lebensmonat und dem 7. Lebensjahr unter anderem das Vorliegen eines Neurodermitisschubs erhoben worden. Die beobachteten Kinder stellen also jetzt die Cluster dar und wir waren daran interessiert, Gruppen von Kindern (bzw. ihren Verläufen) zu finden. Um eine hinreichend flexible Form der Verläufe zu erlauben, waren diese als

Polynome dritter Ordnung in Abhängigkeit vom Alter modelliert worden, jedoch waren datenabhängig auch niedrigere Ordnungen erlaubt. Ausgehend von den Werten des BIC-Kriteriums wurde ein Modell mit 4 Gruppen ausgewählt, in zwei Gruppen waren die Verläufe als konstant, in zwei Gruppen als Polynome dritter Ordnung geschätzt worden. Abbildung 3 zeigt die geschätzten und die mittleren beobachteten Verläufe in diesen Gruppen. Die Gruppe 1 wird von den Kindern gebildet, die praktisch beschwerdefrei bleiben. Die Gruppe 2 zeigt eine altersunabhängige Prävalenz von 23%. In Gruppe 4 finden sich die Kinder, die ab einem gewissen Alter praktisch ständig die Symptome von Neurodermitis zeigen, auch in Gruppe 3 finden wir ein Ansteigen der Neurodermitisprävalenz, allerdings von niedrigerem Niveau aus als bei Gruppe 4.

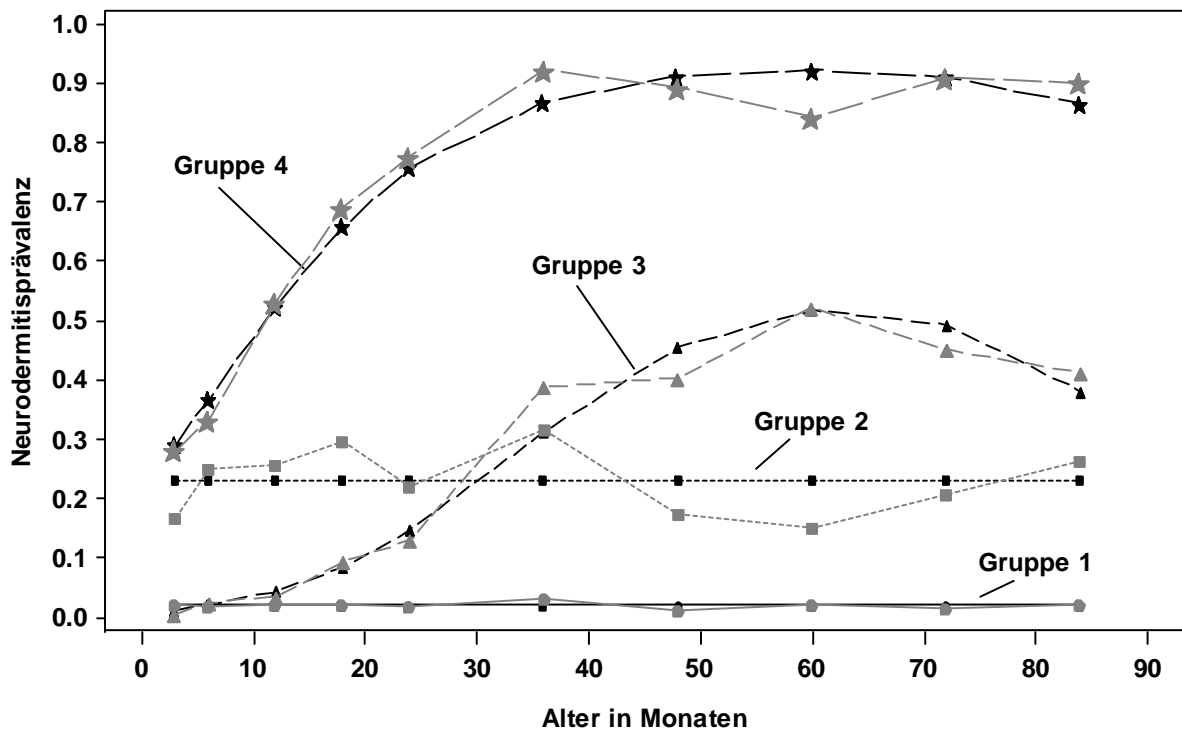


Abbildung 3: Geschätzte und beobachtete Neurodermitisverläufe von 1314 aus der MAS-Studie

Wichtige Kovariablen für die Gruppenzugehörigkeit waren, konsistent mit früheren Analysen, die Familienanamnese mit Neurodermitis und die Länge der Stillzeit¹, wobei eine positive Familienanamnese und eine lange Stillzeit die Prognose verschlechtern.

Das Modell von Nagin verallgemeinernd haben wir zudem vorgeschlagen, die Unsi-

cherheit bei der Kurvenschätzung mit Hilfe von Bootstrap-Konfidenzintervallen sichtbar zu machen. Des Weiteren hat sich die Anwendung des Stereotype-Modells^{110, 111} bei der Modellierung der Kovariablen, die die Gruppenzugehörigkeit beeinflussen, dem herkömmlichen multinomialen logistischen Modell als überlegen erwiesen.

Die Eigenschaften dieses Modells können am besten in den Worten von Nagin selber zusammengefasst werden. Das Modell kann (1) Gruppen von Verläufen identifizieren, (2) den Anteil der Verläufe in den jeweiligen Gruppen schätzen, (3) die Gruppenzugehörigkeit in Abhängigkeit von Kovariablen modellieren und (4) geschätzte durchschnittliche Verläufe in den jeweiligen Gruppen darstellen.

4. Zusammenfassung

Eine Reihe von Studienanlagen in der Medizin führt zu korrelierten oder abhängigen Beobachtungen. Patienten werden z.B. mehrmals im Zeitverlauf, unter verschiedenen experimentellen Bedingungen oder in logischen Einheiten (Clustern) beobachtet. Häufig findet man sogar mehrere dieser Phänomene gleichzeitig und sich überlagernd in einer Studie, also z.B. den Vergleich von Behandlungen an verschiedenen Zeitpunkten in verschiedenen Clustern oder die Erhebung von mehreren Zielgrößen im Zeitverlauf.

Die Standardmethoden der Statistik, die in der Regel eine Unabhängigkeit der Beobachtungen voraussetzen, sind dann nicht mehr gültig. Dies gilt auch für die Gruppe der Regressionsmodelle, deren Ziel es ist, den Einfluss eines oder mehrerer unabhängiger Merkmale auf ein abhängiges Merkmal (Zielgröße) zu beschreiben. Regressionsmodelle mit stetigen Zielgrößen sind auch beim Vorliegen von korrelierten Beobachtungen inzwischen beherrschbar, komplizierter wird die Situation bei nichtstetigen (binären, nominalen oder diskreten) Zielgrößen.

In dieser Arbeit wird zum einen gezeigt, dass Regressionsmodelle für nichtstetige Zielgrößen und korrelierte Beobachtungen in der Medizin immer wichtiger werden. Dieses folgt aus der Analyse der angewandten Methoden im medizinischen Schrifttum im Verlauf der Zeit und aus der Tatsache, dass immer mehr nichttechnische Übersichtsartikel zu diesen Methoden erscheinen.

Zum anderen wird die Relevanz von Regressionsmodellen für nichtstetige Zielgrößen und korrelierte Beobachtungen anhand von vier Anwendungsbeispielen dargestellt. Es werden dabei binäre, nominale und diskrete Zielgrößen betrachtet, außerdem variiert das Anwendungsgebiet von der Kardiochirurgie über die Onkologie und Versorgungsforschung zur Allergologie. Konkret wird gezeigt, dass bei Meta-Analysen für seltene Ereignisse der asymptotische Mantel-Haenszel-Schätzer dem exakten überlegen ist und wie korrelierte Überlebenszeiten aus Tumorregistern ausgewertet werden können. Des Weiteren wird dargestellt, wie Regressionsmodelle für nominale Zielgrößen und korrelierte Beobachtungen in SAS ohne die Anwendung einer Poisson-Likelihood effizienter geschätzt werden können und wie mit Hilfe eines Mischungsmodelles Gruppen von Neurodermitisverläufen datenabhängig identifiziert und dargestellt werden können.

5. Literaturverzeichnis

- (1) Bergmann RL, Diepgen TL, Kuss O et al. Breastfeeding duration is a risk factor for atopic eczema. *Clinical and Experimental Allergy* 2002;32(2):205-9.
- (2) Fluhr JW, Dickel H, Kuss O, Weyher I, Diepgen TL, Berardesca E. Impact of anatomical location on barrier recovery, surface pH and stratum corneum hydration after acute barrier disruption. *British Journal of Dermatology* 2002;146(5):770-6.
- (3) Watzke S, Brieger P, Kuss O, Schoettke H, Wiedl KH. A longitudinal study of learning potential and rehabilitation outcome in schizophrenia. *Psychiatric Services* 2008;59(3):248-55.
- (4) Kuss O, Blankenburg T, Haerting J. A relative survival model for clustered responses. *Biometrical Journal* 2008;50(3):408-18.
- (5) Al-Jundi A, John MT, Setz JM, Szentpetery A, Kuss O. Meta-analytical assessment of treatment need for temporomandibular disorders in non-patient adults. *Journal of Orofacial Pain* 2008;22(2):97-107.
- (6) Parienti JJ, Kuss O. Cluster-crossover design: A method for limiting clusters level effect in community-intervention studies. *Contemporary Clinical Trials* 2007;28(3):316-23.
- (7) Kuss O, Watzke S. [Proper handling of correlated data in rehabilitation research]. *Rehabilitation (Stuttg)* 2005;44(6):367-72.
- (8) Börgermann J, Hakim K, Kuss O et al. Beating versus arrested heart coronary revascularization: Randomized controlled trial in 616 unselected patients (3CAB-STUDY). (*in preparation*) 2008.
- (9) Mccullagh P. Regression-Models for Ordinal Data. *Journal of the Royal Statistical Society Series B-Methodological* 1980;42(2):109-42.
- (10) Nelder JA, Wedderburn RW. Generalized Linear Models. *Journal of the Royal Statistical Society Series A-General* 1972;135(3):370-&.
- (11) Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 1972;34(2):187-&.
- (12) Beitler PJ, Landis JR. A Mixed-Effects Model for Categorical-Data. *Biometrics* 1985;41(4):991-1000.
- (13) Kass RE, Steffey D. Approximate Bayesian-Inference in Conditionally Independent Hierarchical-Models (Parametric Empirical Bayes Models). *Journal of the American Statistical Association* 1989;84(407):717-26.
- (14) Agresti A, Hartzel J. Strategies for comparing treatments on a binary response with multi-centre data. *Statistics in Medicine* 2000;19(8):1115-39.

- (15) Booth JG, Hobert JP. Standard errors of prediction in generalized linear mixed models. *Journal of the American Statistical Association* 1998;93(441):262-72.
- (16) Kuss O. How to Use SAS for Logistic Regression with Correlated Data. *Proceedings of the Twenty-Seventh Annual SAS Users Group International Conference 2002*;P261-27.
- (17) Sartori N, Severini TA. Conditional likelihood inference in generalized linear mixed models. *Statistica Sinica* 2004;14(2):349-60.
- (18) Diggle P, Haegerty P, Liang KY, Zeger S. *Analysis of Longitudinal Data*. 2nd ed. Oxford University Press; 2002.
- (19) Molenberghs G, Verbeke G. *Models for Discrete Longitudinal Data*. Springer; 2005.
- (20) Liang KY, Zeger SL. Longitudinal Data-Analysis Using Generalized Linear-Models. *Biometrika* 1986;73(1):13-22.
- (21) Ren S, Yang S, Lai S. Intraclass correlation coefficients and bootstrap methods of hierarchical binary outcomes. *Statistics in Medicine* 2006;25(20):3576-88.
- (22) Breslow NE, Clayton DG. Approximate Inference in Generalized Linear Mixed Models. *Journal of the American Statistical Association* 1993;88(421):9-25.
- (23) Wolfinger R, Oconnell M. Generalized Linear Mixed Models - A Pseudo-Likelihood Approach. *Journal of Statistical Computation and Simulation* 1993;48(3-4):233-43.
- (24) Crouch EAC, Spiegelman D. The Evaluation of Integrals of the Form $\int_{-\infty}^{\infty} f(t) \exp(-t^2) dt$ - Application to Logistic Normal-Models. *Journal of the American Statistical Association* 1990;85(410):464-9.
- (25) Liu Q, Pierce DA. A Note on Gauss-Hermite Quadrature. *Biometrika* 1994;81(3):624-9.
- (26) Zeger SL, Karim MR. Generalized Linear-Models with Random Effects - A Gibbs Sampling Approach. *Journal of the American Statistical Association* 1991;86(413):79-86.
- (27) Neuhaus JM, Kalbfleisch JD, Hauck WW. A Comparison of Cluster-Specific and Population-Averaged Approaches for Analyzing Correlated Binary Data. *International Statistical Review* 1991;59(1):25-35.
- (28) Zeger SL, Liang KY, Albert PS. Models for Longitudinal Data - A Generalized Estimating Equation Approach. *Biometrics* 1988;44(4):1049-60.
- (29) Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE. *Regression Methods in Biostatistics*. 1st ed. Springer; 2005.

- (30) Horton NJ, Switzer SS. Statistical methods in the journal. *New England Journal of Medicine* 2005;353(18):1977-9.
- (31) Kim M. Statistical methods in Arthritis & Rheumatism: current trends. *Arthritis & Rheumatism* 2006;54(12):3741-9.
- (32) Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. *Journal of the American medical Association* 2007;298(9):1010-22.
- (33) Dendukuri N, Reinhold C. Correlation and regression. *American Journal of Roentgenology* 2005;185(1):3-18.
- (34) Obuchowski NA. Multivariate statistical methods. *American Journal of Roentgenology* 2005;185(2):299-309.
- (35) Crawford SL. Correlation and regression. *Circulation* 2006;114(19):2083-8.
- (36) Griffin D, Audige L. Common statistical methods in orthopaedic clinical studies. *Clinical Orthopaedics and Related Research* 2003;(413):70-9.
- (37) Anderson RP, Jin R, Grunkemeier GL. Understanding logistic regression analysis in clinical reports: an introduction. *Annals of Thoracic Surgery* 2003;75(3):753-7.
- (38) Anderson WN. Statistical techniques for validating logistic regression models. *Annals of Thoracic Surgery* 2005;80(4):1169.
- (39) Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis Part III: multivariate data analysis -- choosing a model and assessing its adequacy and fit. *British Journal of Cancer* 2003;89(4):605-11.
- (40) Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis part II: multivariate data analysis--an introduction to concepts and methods. *British Journal of Cancer* 2003;89(3):431-6.
- (41) Katz MH. Multivariable analysis: a primer for readers of medical research. *Annals of Internal Medicine* 2003;138(8):644-50.
- (42) Zou KH, Tuncali K, Silverman SG. Correlation and simple linear regression. *Radiology* 2003;227(3):617-22.
- (43) Gareen IF, Gatsonis C. Primer on multiple regression models for diagnostic imaging research. *Radiology* 2003;229(2):305-10.
- (44) Bender R, Ziegler A, Lange S. [Logistic regression]. *Deutsche Medizinische Wochenschrift* 2007;132 Suppl 1:e33-e35.
- (45) Bender R, Ziegler A, Lange S. [Multiple regression]. *Deutsche Medizinische Wochenschrift* 2007;132 Suppl 1:e30-e32.
- (46) Lange S, Bender R. [Linear regression and correlation]. *Deutsche Medizinische Wochenschrift* 2007;132 Suppl 1:e9-11.

- (47) Ziegler A, Lange S, Bender R. [Survival analysis: Cox regression]. *Deutsche Medizinische Wochenschrift* 2007;132 Suppl 1:e42-e44.
- (48) Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: Part 2--Correlation between subjects. *British Medical Journal* 1995;310(6980):633.
- (49) Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: Part 1--Correlation within subjects. *British Medical Journal* 1995;310(6977):446.
- (50) Bland JM, Altman DG. Correlation, regression, and repeated data. *British Medical Journal* 1994;308(6933):896.
- (51) Bland JM, Kerry SM. Statistics notes. Trials randomised in clusters. *British Medical Journal* 1997;315(7108):600.
- (52) Kerry SM, Bland JM. Analysis of a trial randomised in clusters. *British Medical Journal* 1998;316(7124):54.
- (53) Vickers AJ, Altman DG. Statistics notes: Analysing controlled trials with baseline and follow up measurements. *British Medical Journal* 2001;323(7321):1123-4.
- (54) Blackstone EH. Breaking down barriers: helpful breakthrough statistical methods you need to understand better. *Journal of Thoracic and Cardiovascular Surgery* 2001;122(3):430-9.
- (55) Bender R, Grouven U, Ziegler A. [Analysis of variance for repeated measurements]. *Deutsche Medizinische Wochenschrift* 2007;132 Suppl 1:e61-e64.
- (56) Kuss O, Gromann C. An exact test for meta-analysis with binary endpoints. *Methods of Information in Medicine* 2007;46(6):662-8.
- (57) Kuss O, McLerran D. A note on the estimation of the multinomial logistic model with correlated responses in SAS. *Computer Methods and Programs in Biomedicine* 2007;87(3):262-9.
- (58) Kuss O, Gromann C, Diepgen TL. Model-based clustering of binary longitudinal atopic dermatitis disease histories by latent class mixture models. *Biometrical Journal* 2006;48(1):105-16.
- (59) Emerson JD. Combining estimates of the odds ratio: the state of the art. *Statistical Methods in Medical Research* 1994;3(2):157-78.
- (60) Agresti A. Exact inference for categorical data: recent advances and continuing controversies. *Statistics in Medicine* 2001;20(17-18):2709-22.
- (61) Mantel N, Haenszel W. Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease. *Journal of the National Cancer Institute* 1959;22(4):719-48.

- (62) Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for Meta-Analysis in Medical Research*. Wiley & Sons, Chichester; 2000.
- (63) Whitehead A. *Meta-analysis of controlled clinical trials*. Wiley & Sons, Chichester; 2002.
- (64) Cheng DC, Bainbridge D, Martin JE, Novick RJ. Does off-pump coronary artery bypass reduce mortality, morbidity, and resource utilization when compared with conventional coronary artery bypass? A meta-analysis of randomized trials. *Anesthesiology* 2005;102(1):188-203.
- (65) Kuss O, Börgermann J, Gummert JF. Meta-analyses with rare events should use adequate methods. *Journal of Thoracic and Cardiovascular Surgery* 2008;136(1):241.
- (66) Gondos A, Holleczeck B, Arndt V, Stegmaier C, Ziegler H, Brenner H. Trends in population-based cancer survival in Germany: to what extent does progress reach older patients? *Annals of Oncology* 2007;18(7):1253-9.
- (67) Buckley JD. Additive and multiplicative models for relative survival rates. *Biometrics* 1984;40(1):51-62.
- (68) Esteve J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: elements for further discussion. *Statistics in Medicine* 1990;9(5):529-38.
- (69) Berkson J, Gage RP. Calculation of survival rates for cancer. *Proc Staff Meet Mayo Clin* 1950;25(11):270-86.
- (70) Hakulinen T, Tenkanen L. Regression-Analysis of Relative Survival Rates. *Applied Statistics-Journal of the Royal Statistical Society Series C-Applied Statistics* 1987;36(3):309-17.
- (71) Weller EA, Feuer EJ, Frey CM, Wesley MN. Parametric relative survival regression using generalized linear models with application to Hodgkin's lymphoma. *Journal of the Royal Statistical Society Series C-Applied Statistics* 1999;48:79-89.
- (72) Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Statistics in Medicine* 2004;23(1):51-64.
- (73) Bollmann A, Blankenburg T, Haerting J et al. Survival of patients in clinical stages I-IIIb of non-small-cell lung cancer treated with radiation therapy alone - Results of a population-based study in Southern Saxony-Anhalt. *Strahlentherapie und Onkologie* 2004;180(8):488-96.
- (74) Prentice RL, Zhao LP. Estimating Equations for Parameters in Means and Covariances of Multivariate Discrete and Continuous Responses. *Biometrics* 1991;47(3):825-39.
- (75) Vonesh EF, Wang H, Nie L, Majumdar D. Conditional second-order generalized estimating equations for generalized linear and nonlinear mixed-effects models. *Journal of the American Statistical Association* 2002;97(457):271-83.

- (76) Hall DB, Severini TA. Extended generalized estimating equations for clustered data. *Journal of the American Statistical Association* 1998;93(444):1365-75.
- (77) Preisser JS, Qaqish BF. Robust regression for clustered data with application to binary responses. *Biometrics* 1999;55(2):574-9.
- (78) Stoner JA, Leroux BG. Analysis of clustered data: A combined estimating equations approach. *Biometrika* 2002;89(3):567-78.
- (79) Sutradhar BC. An overview on regression models for discrete longitudinal responses. *Statistical Science* 2003;18(3):377-93.
- (80) Chiou JM, Muller HG. Estimated estimating equations: semiparametric inference for clustered and longitudinal data. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 2005;67:531-53.
- (81) Qu A, Lindsay BG, Li B. Improving generalised estimating equations using quadratic inference functions. *Biometrika* 2000;87(4):823-36.
- (82) Goldstein H, Rasbash J. Improved approximations for multilevel models with binary responses. *Journal of the Royal Statistical Society Series A-Statistics in Society* 1996;159:505-13.
- (83) McCulloch CE. Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American Statistical Association* 1997;92(437):162-70.
- (84) Jiang J. Consistent estimators in generalized linear mixed models. *Journal of the American Statistical Association* 1998;93(442):720-9.
- (85) Lee Y, Nelder JA. Hierarchical generalized linear models. *Journal of the Royal Statistical Society Series B-Methodological* 1996;58(4):619-56.
- (86) Lee Y, Nelder JA. Double hierarchical generalized linear models. *Journal of the Royal Statistical Society Series C-Applied Statistics* 2006;55:139-67.
- (87) Zhu HT, Lee SY. Analysis of generalized linear mixed models via a stochastic approximation algorithm with Markov chain Monte-Carlo method. *Statistics and Computing* 2002;12(2):175-83.
- (88) Kuhn E, Lavielle M. Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics & Data Analysis* 2005;49(4):1020-38.
- (89) Pan JX, Thompson R. Quasi-Monte Carlo estimation in generalized linear mixed models. *Computational Statistics & Data Analysis* 2007;51(12):5765-75.
- (90) Pan JX, Thompson R. Gauss-Hermite quadrature approximation for estimation in generalised linear mixed models. *Computational Statistics* 2003;18(1):57-78.
- (91) Aitkin M. A general maximum likelihood analysis of variance components in generalized linear models. *Biometrics* 1999;55(1):117-28.

- (92) Mills JE, Field CA, Dupuis DJ. Marginally specified generalized linear mixed models: A robust approach. *Biometrics* 2002;58(4):727-34.
- (93) Bellio R, Varin C. A pairwise likelihood approach to generalized linear models with crossed random effects. *Statistical Modelling* 2005;5(3):217-27.
- (94) Heagerty PJ, Zeger SL. Marginalized multilevel models and likelihood inference. *Statistical Science* 2000;15(1):1-19.
- (95) Karcher P, Wang YD. Generalized nonparametric mixed effects models. *Journal of Computational and Graphical Statistics* 2001;10(4):641-55.
- (96) Goldstein H. Restricted Unbiased Iterative Generalized Least-Squares Estimation. *Biometrika* 1989;76(3):622-3.
- (97) Pinheiro JC, Bates DM. Approximations to the Log-likelihood Function in the Nonlinear Mixed-Effects Model. *Journal of Computational and Graphics Statistics* 1995;4(1):12-35.
- (98) Raudenbush SW, Yang ML, Yosef M. Maximum Likelihood for Generalized Linear Models With Nested Random Effects via High-Order, Multivariate Laplace Approximation. *Journal of Computational and Graphical Statistics* 2000;9(1):141-57.
- (99) Sinha SK. Robust analysis of generalized linear mixed models. *Journal of the American Statistical Association* 2004;99(466):451-60.
- (100) Vaida F, Meng XL. Two slice-EM algorithms for fitting generalized linear mixed models with binary response. *Statistical Modelling* 2005;5(3):229-42.
- (101) Vonesh EF, Wang H, Majumdar D. Generalized least squares, Taylor series linearization, and Fisher's scoring in multivariate nonlinear regression. *Journal of the American Statistical Association* 2001;96(453):282-91.
- (102) Vonesh EF, Chinchilli VM. *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. New York: Marcel Dekker; 1997.
- (103) Yau KKW, Kuk AYC. Robust estimation in generalized linear mixed models. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 2002;64:101-17.
- (104) Hartzel J, Agresti A, Caffo B. Multinomial logit random effects models. *Statistical Modelling* 2001;1(2):81-102.
- (105) Chen Z, Kuo L. A note on the estimation of the multinomial logit model with random effects. *American Statistician* 2001;55(2):89-95.
- (106) Malchow-Moller N, Svarer M. Estimation of the multinomial logit model with random effects. *Applied Economics Letters* 2003;10(7):389-92.
- (107) Mccullagh P, Nelder JA. *Generalized Linear Models*. London: Chapman & Hall; 1989.

- (108) Wright SP. Multivariate analysis using the MIXED procedure. *Proceedings of the 23rd Annual SAS Users Group (SUGI) International Conference* 1998;P229-23.
- (109) Hasenbein U, Kuss O, Baumer M, Schert C, Schneider H, Wallesch CW. Physicians' preferences and expectations in traumatic brain injury rehabilitation - results of a case-based questionnaire survey. *Disability and Rehabilitation* 2003;25(3):136-42.
- (110) Anderson JA. Regression and Ordered Categorical Variables. *Journal of the Royal Statistical Society Series B-Methodological* 1984;46(1):1-30.
- (111) Kuss O. On the estimation of the stereotype regression model. *Computational Statistics & Data Analysis* 2006;50(8):1877-90.
- (112) Kuss O. Modelling physicians' recommendations for optimal medical care by random effects stereotype regression. *Proceedings of the 18th International Workshop on Statistical Modelling; Leuven: Katholieke Universiteit*, 2003 p. 245-9.
- (113) Nagin DS. Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods* 1999;4(2):139-57.
- (114) Bergmann RL, Bergmann KE, Lau-Schadensdorf S et al. Atopic diseases in infancy. The German multicenter atopy study (MAS-90). *Pediatric Allergy and Immunology* 1994;5(6 Suppl):19-25.

6. Anlagen

Kuss O, Gromann C. An exact test for meta-analysis with binary endpoints. Methods of Information in Medicine 2007;46(6):662-8.

Kuss O, Blankenburg T, Haerting J. A relative survival model for clustered responses. Biometrical Journal 2008;50(3):408-18.

Kuss O, McLerran D. A note on the estimation of the multinomial logistic model with correlated responses in SAS. Computer Methods and Programs in Biomedicine 2007;87(3):262-9.

Kuss O, Gromann C, Diepgen TL. Model-based clustering of binary longitudinal atopic dermatitis disease histories by latent class mixture models. Biometrical Journal 2006;48(1):105-16.

7. Thesen

Eine Reihe von Studienanlagen in der Medizin führt zu korrelierten oder abhängigen Beobachtungen. Patienten werden z.B. mehrmals im Zeitverlauf, unter verschiedenen experimentellen Bedingungen oder in logischen Einheiten (Clustern) beobachtet. Häufig findet man sogar mehrere dieser Phänomene gleichzeitig und sich überlagernd in einer Studie. Die Standardmethoden der Statistik, die in der Regel eine Unabhängigkeit der Beobachtungen voraussetzen, sind dann nicht mehr gültig.

Dies gilt auch und insbesondere für die Gruppe der Regressionsmodelle, deren Ziel es ist, den Einfluss eines oder mehrerer unabhängiger Merkmale auf ein abhängiges Merkmal (Zielgröße) zu beschreiben.

Nicht immer (oder besser: eher selten) hat man es in der Medizin mit stetigen Zielgrößen zu tun. Viel häufiger treten nichtstetige Zielgrößen auf, unter denen die binären Zielgrößen (z.B. das Vorliegen einer Krankheit oder das Eintreten eines unerwünschten Ereignisses) die prominenteste Gruppe darstellt. Des Weiteren treten ordinale, nominale und diskrete Zielgrößen auf.

Regressionsmodelle mit stetigen Zielgrößen sind auch beim Vorliegen von korrelierten Beobachtungen inzwischen beherrschbar, komplizierter wird die Situation bei nichtstetigen (binären, nominalen oder diskreten) Zielgrößen.

Regressionsmodelle für nichtstetige Zielgrößen und korrelierte Beobachtungen werden in der Medizin immer häufiger. Diese kann man zum einen sehen, wenn man die angewandten Methoden im medizinischen Schrifttum im Verlauf der Zeit auswertet, zum anderen daran, dass immer mehr nichttechnische Übersichtsartikel zu diesen Methoden erscheinen.

Die Entwicklung von statistischen Methoden für Regressionsmodelle für nichtstetige Zielgrößen und korrelierte Beobachtungen eilt der tatsächlichen Anwendung der Methoden weit voraus. Hier herrscht noch einiger Nachholbedarf, vor allem was die Umsetzung der entwickelten Methoden in gängige Software angeht.

Das Anwendungsspektrum für Regressionsmodelle für nichtstetige Zielgrößen und korrelierte Beobachtungen in der Medizin ist weit. Die hier vorgestellten Beispiele aus der Kardiochirurgie, der Onkologie, der Versorgungsforschung und der Allergologie stellen nur einen kleinen Ausschnitt aus diesem Spektrum dar.

In Meta-Analysen für seltene Ereignisse ist der asymptotische Mantel-Haenszel-Schätzer dem exakten überlegen. Die Standardmethoden (Modell mit festen und mit zufälligen Effekten) sollten bei seltenen Ereignissen generell nicht eingesetzt werden.

Ein Modell für die relative Überlebenszeitanalyse für korrelierte Beobachtungen kann als ein Spezialfall der Klasse der Generalisierten Linearen Gemischten Modelle interpretiert und geschätzt werden. Bei der Anwendung in einer versorgungsepidemiologischen Studie erweist sich das Modell als den Standardmethoden überlegen.

Regressionsmodelle für nominale Zielgrößen und korrelierte Beobachtungen können ohne den Rückgriff auf die Poisson-Likelihoodfunktion effizienter geschätzt werden. Die Statistik-Software SAS erlaubt dabei die Schätzung sowohl marginaler als auch gemischter Modelle.

Mit Hilfe eines Mischungsmodells für latente Gruppen können Neurodermitisverläufen datenabhängig identifiziert und dargestellt werden. Die Unsicherheit bei der Kurvenschätzung kann dabei mit Hilfe von Bootstrap-Konfidenzintervallen sichtbar gemacht werden. Bei der Modellierung von Kovariablen, die die Gruppenzugehörigkeit beeinflussen, kann neben dem herkömmlichen multinomialen logistischen Modell das Stereotype-Modell eingesetzt werden.

8. Tabellarischer Lebenslauf

Name Dr. Kuß
Vorname Oliver
Geburtsort Crailsheim, Baden-Württemberg
Geburtsdatum 16.07.1969



Adresse (dienstlich) Institut für Medizinische Epidemiologie, Biometrie und Informatik
Martin-Luther-Universität Halle-Wittenberg
Magdeburger Str. 8
06097 Halle (Saale)
Tel.: 0345/557-3582
Fax: 0345/557-3580
Email: oliver.kuss@medizin.uni-halle.de
WWW: www.oliverkuss.de/science/

Adresse (privat) Goethestr. 34
06114 Halle (Saale)
Tel.: 0345/6784430
Tel. (mobil): 0174/3101216

Schulischer Werdegang August 1975 – Grundschole Fichtenau
Mai 1979

August 1979 – Albert-Schweitzer-Gymnasium Crailsheim
Mai 1988

Erfüllung der Wehrpflicht August 1988 – Ableistung des Zivildienstes bei der Arbeiterwohlfahrt, Kreisverband Schwäbisch Hall
Februar 1990

Universitärer Werdegang April 1990 – September 1991 Studium der Mathematik und der Politischen Wissenschaften an der Universität Heidelberg

Oktober 1991 – Studium der Mathematik und der Geographie an der Universität Heidelberg
November 1996

30.11.1996 Staatsexamen für das Lehramt an Gymnasien

Beruflicher Werdegang	April 1997 – Dezember 1999	Wissenschaftlicher Angestellter (BAT IIa/2), Dermatologische Universitätsklinik Erlangen
	Januar 2000 – Dezember 2000	Wissenschaftlicher Angestellter (BAT IIa/2), Abteilung Klinische Sozialmedizin, Universi- tätsklinikum Heidelberg
	Januar 2001 – heute	Wissenschaftlicher Angestellter (BAT IIa-O), Institut für Medizinische Epidemiologie, Bi- ometrie und Informatik, Universität Halle- Wittenberg
	24.03.2003	Promotion zum Dr. sc. hum. an der Medizi- nischen Fakultät der Universität Heidelberg (Note: magna cum laude)