

Charité - Universitätsmedizin Berlin
Campus Benjamin Franklin
Aus dem Institut für Sozialmedizin, Epidemiologie und
Gesundheitsökonomie
Direktor: Prof. Dr. Stefan N. Willich
Abteilung: Internationale Gesundheitswissenschaften
Abteilungsleiter: Prof. Dr. Dankmar Böhning

Untersuchung von verschiedenen
Modellierungen der Heterogenität in
multizentrischen Studien

Inaugural – Dissertation
zur Erlangung des Grades
Doctor rerum medicarum
der Charité – Universitätsmedizin Berlin
Campus Benjamin Franklin

vorgelegt von Ronny Kuhnert
aus Löbau

Referent: Prof. Dr. Dankmar Böhning

Koreferent: Prof. Dr. Peter Martus

Gedruckt mit der Genehmigung der Charité – Universitätsmedizin Berlin
Campus Benjamin Franklin

Promoviert am: 19.07.2005

Inhaltsverzeichnis

Einführung	ix
1 Multizentrische Studien	1
2 Die drei Modellansätze	17
2.1 Allgemeine Modellierung der Heterogenität	17
2.2 Normalverteilungsansatz	19
2.3 Profil-Likelihoodansatz	22
2.4 Multi-Level-Modell	23
3 Schätzen der Parameter	27
3.1 Parameterschätzung im Normalverteilungs- Modell	28
3.2 Parameterschätzung im Profil-Likelihood- Ansatz	28
3.3 Parameterschätzung im Multi-Level-Modell	29
3.4 EM-Algorithmus bei einer Mischverteilung	31
3.5 Eigenschaften von Mischverteilungen	32
3.5.1 Gradientenfunktion	33
3.5.2 Der allgemeine Maximum-Likelihood Satz von Mischverteilungen	34
3.6 EM-Algorithmus mit Gradientenfunktion	36
3.6.1 EM-Algorithmus mit Gradientenfunktions-Update bei bekannter An- zahl der Komponenten	36
3.6.2 EM-Algorithmus mit Gradientenfunktions-Update bei unbekannter Anzahl der Komponenten (gEMGFU)	37
3.7 gEMGFU in FORTRAN	38
3.7.1 Numerische Stabilität	39
3.7.2 Die Maximierung der Gradientenfunktion	43

4	Auswertung der Studien	47
4.1	Ohne Heterogenität (Homogenitätsfall)	47
4.2	Mit unbeobachteter Heterogenität	47
4.2.1	Beta-Blocker Studie	48
4.2.2	Studie über Infektion der Atemwege	49
4.2.3	Studie über Präeklampsie	51
4.2.4	CALGB-Studie	52
4.2.5	Multi-Farm-Studie	52
4.2.6	Herzinfarkt-Studie	52
4.2.7	Studie über Frühgeburten	62
4.2.8	Studie über Bluthochdruck	62
4.2.9	Studie über koronare Herzkrankheit	63
4.2.10	Studie über Vitamin A Ergänzung	65
4.2.11	Brustkrebsvorsorgestudie	75
4.2.12	Nikotinersatzstudie	75
4.2.13	Kaiserschnittsstudie	76
4.2.14	Zusammenfassung	83
5	Simulationsstudie	85
5.1	Unter Effekthomogenität	85
5.1.1	Beta-Blocker Studie	86
5.1.2	CALGB-Studie	87
5.1.3	Multi-Farm-Studie	87
5.1.4	Studie über koronare Herzkrankheit	90
5.2	Unter Effektheterogenität	93
5.2.1	Beta-Blocker Studie	94
5.2.2	CALGB-Studie	97
5.2.3	Studie über koronare Herzkrankheit	98
5.2.4	Eigenständige Simulationsstrategien	104
5.3	Schätzung der Komponentenzahl	111
5.3.1	Beta-Blocker Studie	112
5.3.2	Eigenständige Simulationsstudie	113
5.4	Stabilität der Simulationsstudien	115
6	Diskussion	117
6.1	Zusammenfassung	120

A	123
A.1 Bootstrap-Stichproben	123
A.2 Ergebnisse einzelner Bootstrap-Stichproben	126
A.3 Ableitungen der Gradientenfunktion	133
Literaturverzeichnis	137
Danksagung	141

Abbildungsverzeichnis

2.1	Empirische Verteilung des log Relativen Risikos im Vergleich mit der Normalverteilung (Daten sind aus Tabelle 1.12)	21
3.1	CALGB-Studie (Bias vom Relativen Risiko-Schätzer wurde an 30 Stellen gleichverteilt im Intervall [0.01;0.99] berechnet, die generierten x^T , x^C -Werte sind binomialverteilt, additive Konstante ist 0.00000001)	42
3.2	CALGB-Studie (Bias vom Relativen Risiko-Schätzer wurde an 30 Stellen gleichverteilt im Intervall [0.01;0.99] berechnet, die generierten x^T , x^C -Werte sind binomialverteilt, additive Konstante ist 0.5)	43
4.1	Zentrenzugehörigkeit bei der Atemwegsstudie mit zwei Mischungskomponenten (Profil-Likelihood-Ansatz). Die Zahlen neben den Punkten geben das Zentrum an.	50
4.2	Zentrenzugehörigkeit bei der Atemwegsstudie mit zwei Mischungskomponenten (Normalverteilungsansatz). Die Zahlen neben den Punkten geben das Zentrum an.	51
4.3	Zentrenzugehörigkeit bei der Atemwegsstudie mit vier Mischungskomponenten (Multi-Level-Modell). Die Zahlen neben den Punkten geben das Zentrum an.	54
4.4	Zentrenzugehörigkeit bei der Studie über koronarer Herzkrankheit mit vier Mischungskomponenten (Profil-Likelihood-Ansatz). Die Zahlen neben den Punkten geben das Zentrum an.	64
4.5	Zentrenzugehörigkeit bei der Studie über koronarer Herzkrankheit mit vier Mischungskomponenten (Normalverteilungsansatz). Die Zahlen neben den Punkten geben das Zentrum an.	64

5.1	Beta-Blocker Studie (Bias, Varianz und MSE vom Relativen Risiko-Schätzer wurde an 30 Stellen gleichverteilt im Intervall [0.01;4.73] berechnet, die generierten x^T , x^C - Werte sind binomialverteilt)	88
5.2	CALGB-Studie (Bias, Varianz und MSE vom Relativen Risiko-Schätzer wurde an 30 Stellen gleichverteilt im Intervall [0.01;0.99] berechnet, die generierten x^T , x^C - Werte sind binomialverteilt)	89
5.3	Multi-Farm-Studie (Bias, Varianz und MSE vom Relativen Risiko-Schätzer wurde an 30 Stellen gleichverteilt im Intervall [0.01;5.99] berechnet, die generierten x^T , x^C - Werte sind binomialverteilt)	91
5.4	Studie über koronarer Herzkrankheit (Bias, Varianz und MSE vom Relativen Risiko-Schätzer wurde an 30 Stellen gleichverteilt im Intervall [0.01;6.85] berechnet, die generierten x^T , x^C - Werte sind Poissonverteilt)	92
5.5	Histogramm des Relativen Risikos für jede Komponente (Beta-Blocker Studie).	95
5.6	Histogramm von den Gewichten für jede Komponente (Beta-Blocker Studie).	96
5.7	Histogramm des Relativen Risikos für drei Komponenten (Beta-Blocker Studie).	98
5.8	Histogramm mit den drei Komponenten zusammen, je nach Modell (Beta-Blocker Studie)	99
5.9	Stabilität der Simulation im Normalfall (Tabelle 5.14, $n \sim Poi(3000)$). Abgebildet ist der Erwartungswert der jeweiligen Komponente für das Relative Risiko in Abhängigkeit der Simulationswiederholungen.	108
5.10	Stabilität der Simulation im Normalfall (Tabelle 5.10). Abgebildet ist der Erwartungswert der jeweiligen Komponente für das Relative Risiko in Abhängigkeit der Simulationswiederholungen.	116