

Berufsbedingte Atemwegserkrankungen, Abschließender Praktikumsbericht

Ursula Gerhardinger, Florian Hiemeyer,
Axel Kraus, Anja Kühnlein

30. Januar 2004

Struktur des Vortrags:

- Aufbau/ Ziele der verwendeten Studie
- Erhobene Variablen, Datenstruktur
- Scorebildung
- Deskriptive Datenauswertung
- Non-Response-Analyse
- Modellwahl

- Umkodierungen
- Der Weg zum Endmodell
- Das Endmodell
- Interpretation
- Einfache longitudinale Analyse
- Survivalanalyse
- Fazit

Aufbau und Ziele

- Erhebungen in Hamburg und Erfurt

	Hamburg	Erfurt
1991/1992	1259	732
2001/2002	303	287

- Bei der 2. Erhebung: Zusätzlich *retrospektive* Erhebung der Berufe der vergangenen zehn Jahre; Codierung der Berufe zu Expositionswerten
- **Ziel:** Lassen sich Erkrankungen durch Schadstoffexposition erklären?

Datenstruktur I

Beschränkung auf einige demographische Daten, vier Expositionstypen, und zwei Krankheitstypen

Erkrankungen:

- Rhinitis (chronischer Schnupfen)
- allergischer Schnupfen

Datenstruktur II

demographische Daten/Prädiktoren/Confounder:

- Wohnort (Studienzentrum)
- Geschlecht
- Alter
- Bildungsniveau
- Lebensumstände im Kleinkindalter

demographische Daten/Prädiktoren/Confounder:

- Allergie bzw. Asthma in der Familie (genetische Vorbelastung)
- Rauchverhalten

Datenstruktur III

Exposition

Aus der Berufshistorie \Rightarrow Berechnung von Expositionen am Arbeitsplatz.

- biologische Exposition: durch Viren, Bakterien, sonstige biologische Luftbelastungen
- mineralische Exposition: durch Stäube von Kohle, und Metallspänen, etc.
- Gas/Rauch Exposition: durch feinste Feststoffpartikel und durch Gase

Datenstruktur III

Dazu kommt die Passivrauchexposition, die allerdings nur zu zwei Zeitpunkten (erste und zweite Erhebung), und nicht in zeitlicher Dauer erfasst wurde.

Exposition gegenüber Passivrauch:

nie	→	0
bei der ersten Untersuchung	→	1
bei der zweiten Untersuchung	→	2
bei beiden Untersuchungen	→	3

Scorebildung

Beschäftigungen (B_1, \dots, B_n) mit Schadstoffbelastung (S_1, \dots, S_n) ;

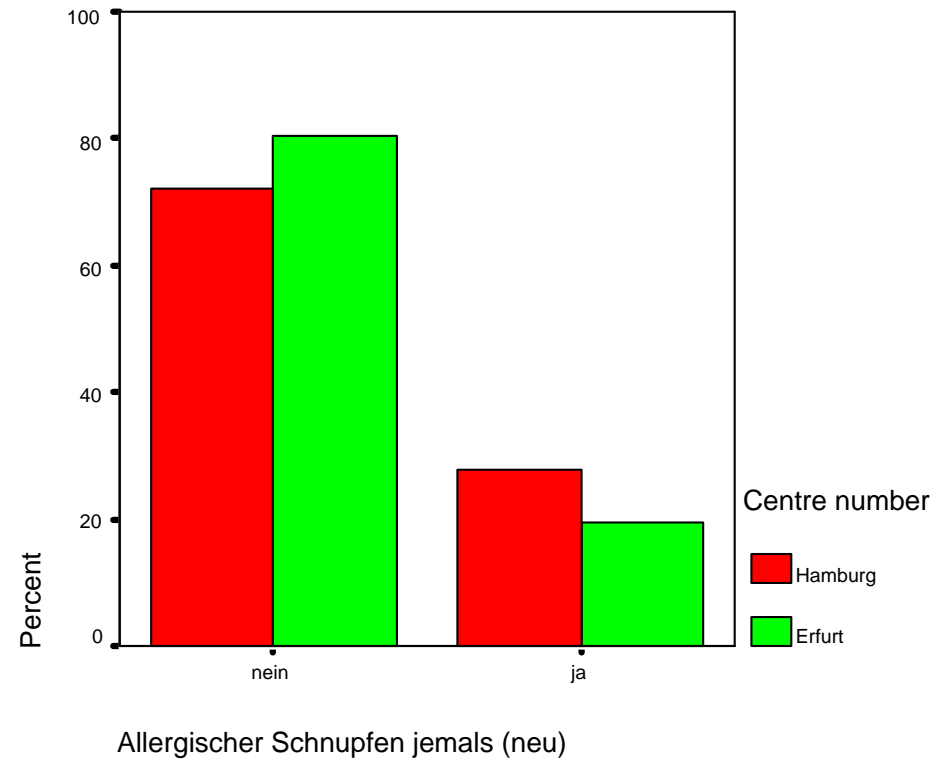
zu jedem $B_i \quad \exists \quad (\Delta t)_i$ für $i = 1, \dots, n$

$$b_{gesamt} = \sum_{i=1}^n S_i * (\Delta t)_i$$

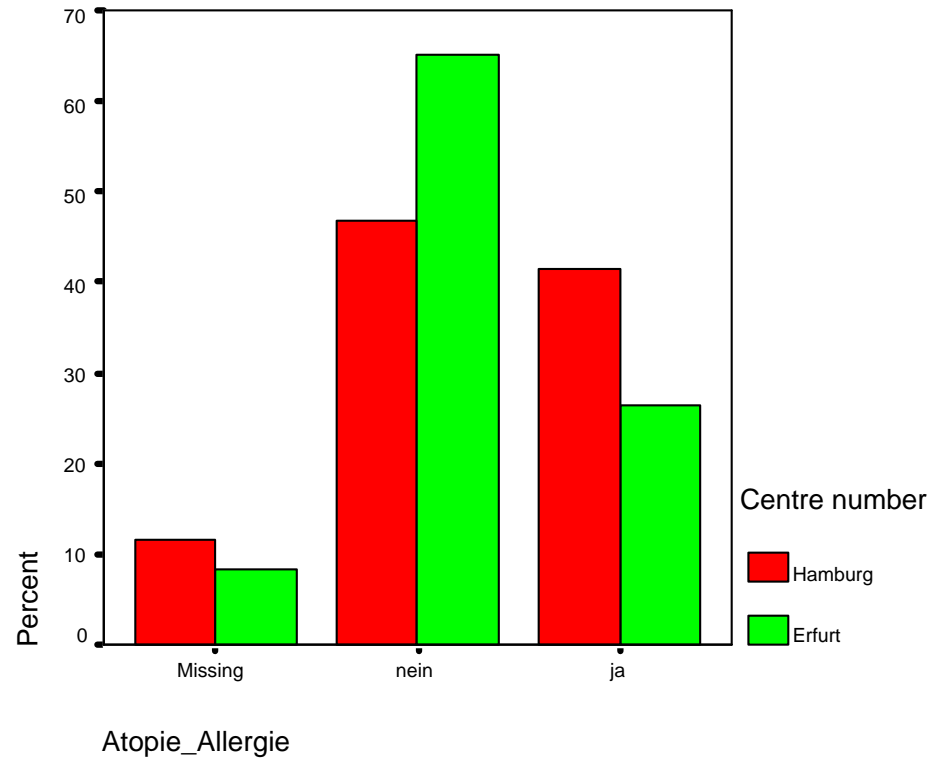
Die Berechnung für m_{gesamt} und g_{gesamt} verläuft nach dem gleichem Schema.

Deskriptive Datenauswertung

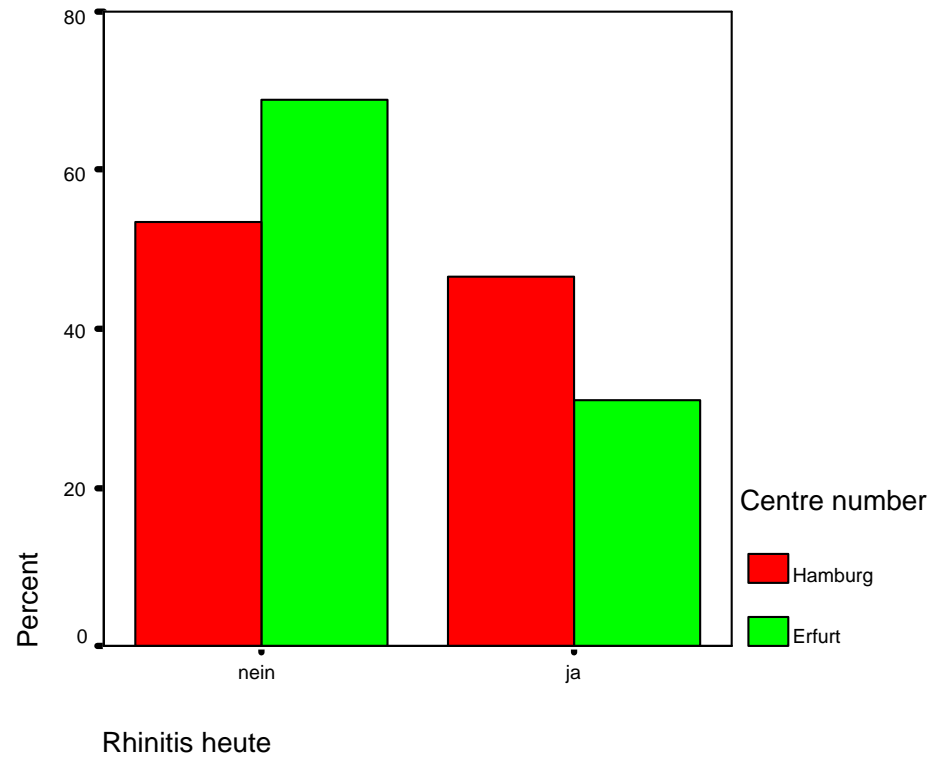
Allergie versus Zentrum



Allergie in der Familie versus Zentrum



Rhinitis versus Zentrum



Korrelationen der Expositionen



Korrelationswerte

Variable	BGesamt	MGesamt	GGesamt
BGesamt	1	0.343	0.383
MGesamt	0.343	1	0.692
GGesamt	0.383	0.692	1

Tabelle 1: Korrelationstabelle der Gesamtexpositionen, alle signifikant zum Niveau $\alpha = 0.01$

Non-Response Analyse

	Erfurt	Hamburg	Erfurt und Hamburg
Anzahl Probanden			
response (= 1. und 2. Untersuchung)	287	303	590
nonresponse (= nur 1. Untersuchung)	445	956	1401
alle	732	1259	1991
response (in %)	39.2	24.1	29.6

Tabelle 2: Übersicht über das Antwortverhalten

	Erfurt	Hamburg
in %		
response	58.9	67.7
nonresponse	67.6	67.9

Tabelle 3: Verteilung der Raucher

Ergebnis:

- Signifikant unterschiedliche Verteilung beim *Rauchen* in Erfurt

Modellwahl

Startmodell:

- logistisches Modell
- Zielgröße: Rhinitis und allerg. Schnupfen
- erklärende Variablen: Exposition, Studienzentrum, Confounder, Interaktionen

Umkodierungen I

Variable	Referenzkategorie
Geschlecht	Männer
Bildungsstand	niedrige Bildung
Wohnort bis 5. LJ	städtische Umgebung
Rauchen	Nie geraucht
Passivrauchen	kein Passivrauch
Allergie in der Familie	keine Vorbelastung
Asthma in der Familie	keine Vorbelastung

Tabelle 4: Dummykodierung der Variablen

Umkodierungen II

Schadstoffexpositionen: extrem rechtsschief verteilt

⇒ Einteilung in Klassen:

Kategorie	0	1	2	3
	Exposition = 0	x_{33}	x_{66}	x_{100}
bgesamt_t	86.95%	4.07%	4.57%	4.41%
mgesamt_t	74.58%	8.47%	8.48%	8.47%
ggesamt_t	73.90%	8.81%	8.82%	8.47%

Ausgangsmodell

Es wird ein logistisches Modell, also

$$\text{logit}[\pi(\mathbf{x}_i)] = \mathbf{x}_i' \boldsymbol{\beta}, \quad i = 1, \dots, n,$$

geschätzt. In den \mathbf{x}_i enthalten sind dabei der Intercept und

- Schadstoff-Expositionen (kategorisiert)
- Confounder
- Interaktionen zwischen
 - Zentrum und Expositionen
 - Geschlecht und Expositionen

Modellverfeinerung

- Interaktionen nicht signifikant \Rightarrow weggelassen
- Confounderanalyse für Wohnort bis zum 5. Lebensjahr
- Effektmodifizierer Zentrum und Geschlecht

Confounderanalyse

Rhinitis-Modell		
mit Ort	nur Intercept	Intercept und Kovariaten
AIC	680.087	685.181
ohne Ort	nur Intercept	Intercept und Kovariaten
AIC	680.087	682.458
Allergie-Modell		
mit Ort	nur Intercept	Intercept und Kovariaten
AIC	571.600	543.829
ohne Ort	nur Intercept	Intercept und Kovariaten
AIC	571.600	546.176

⇒ **allerg. Schnupfen:** Modell mit Ortsvariablen besser

⇒ beide Modelle mit Ortsvariablen wegen besserer Vergleichbarkeit

Effektmodifizierer

Betrachtete Effektmodifizierer: „Zentrum“ und „Geschlecht“

Probleme:

- zu kleine Fallzahlen
- riesige Konfidenzintervalle

⇒ Stratifizierung jeweils einzeln nach „Zentrum“ und „Geschlecht“

Endmodelle für Rhinitis und allergischen Schnupfen

- Keine Interaktionen
- bei Rhinitis: Betrachtung jeweils stratifiziert nach Zentrum und Geschlecht

Unabhängige Variablen:

- biologische-, mineralische- und Gas/Rauch-Exposition
- Alter
- Geschlecht
- Bildungsstand

- Rauchen
- Passivrauchen
- Wohnort bis zum fünften Lebensjahr
- Allergie in der Familie
- Asthma in der Familie

Effekt	Punktschätzer	untere KI-Grenze	obere KI-Grenze
BDUMMYU	1.271	0.405	3.986
BDUMMYM	1.077	0.414	2.804
BDUMMYO	0.670	0.247	1.820
ELO3ALT	0.980	0.953	1.007
SEXDUMMY	1.130	0.753	1.695
ZENTRUM	0.636	0.433	0.935
RAUDUMMY	1.030	0.691	1.535
ATOP_ALL	2.227	1.502	3.301
ATOP_AST	1.023	0.580	1.802
ORTDUM1	0.759	0.439	1.311
ORTDUM2	0.773	0.357	1.674

Tabelle 5: Odds Ratios für Rhinitis

Effekt	Punktschätzer	untere KI-Grenze	obere KI-Grenze
BDUMMYU	1.333	0.310	5.727
BDUMMYM	1.525	0.486	4.783
BDUMMYO	0.859	0.246	2.998
ELO3ALT	0.981	0.950	1.014
SEXDUMMY	1.313	0.811	2.128
ZENTRUM	0.794	0.488	1.291
RAUDUMMY	0.637	0.402	1.008
ATOP_ALL	4.136	2.620	6.530
ATOP_AST	0.722	0.368	1.415
ORTDUM1	0.640	0.328	1.249
ORTDUM2	0.273	0.077	0.973

Tabelle 6: Odds Ratios für allergischen Schnupfen

Interpretation: Rhinitis-Gesamtmodell

- vorhandene Allergie in der Familie steigert Erkrankungsrisiko
- Erfurter haben niedrigeres Krankheitsrisiko als Hamburger

Interpretation: allerg. Schnupfen-Gesamtmodell

- vorhandene Allergie in der Familie steigert Erkrankungsrisiko \Rightarrow stärker als bei Rhinitis
- Wohnort bis zum 5. Lebensjahr auf dem Bauernhof \Rightarrow niedrigeres Risiko
- Nie-Raucher \Rightarrow niedrigeres Risiko

Interpretation: Zentrum-stratifiziertes Rhinitis-Modell

Hamburg

- vorhandene Allergie in der Familie steigert Erkrankungsrisiko
- höhere biologische Exposition „senkt“ Risiko \Rightarrow „Healthy Worker Effect“

Erfurt

- vorhandene Allergie in der Familie steigert Erkrankungsrisiko
- *kein* „Healthy Worker Effect“

Einfache longitudinale Betrachtung

- longitudinale Betrachtung nur für allergischen Schnupfen
- 'Kürzung' der Expositionsvariablen (nur bis zum Zeitpunkt der Erkrankung kumuliert)
- Berechnung des Allergie-Endmodells mit gekürzten Expositionsvariablen
- keine Ergebnisse, da Modell nicht konvergiert

Exkurs zur Survivalanalyse

- **wiederholte** Betrachtung von Probanden **bis zum Erkrankungszeitpunkt**
- registriert wird hier die **Dauer** bis zum Auftreten der Krankheit
- Zielvariable informiert über **Zensierungsstatus**
- hier: **diskrete** Zeitachse

Komplexe longitudinale Betrachtung

- Survivalanalyse
- links und rechts zensiert
- zehn Zeitpunkte
- Expositionsscores: Belastung im jeweils aktuellen Jahr
- **zeitabhängig**: Expositionsscores, Alter und Jahr
- **zeitunabhängig**: Zentrum, Geschlecht, Bildungsstand, Passivrauchen, Rauchen, Allergie in der Familie, Asthma in der Familie, Wohnort

Effekt	Punktschätzer	untere KI-Grenze	obere KI-Grenze
Jahr 9 vs 0	0.109	0.012	0.948
Jahr 1 vs 0	2.030	0.529	7.798
baktuell	0.944	0.818	1.091
maktuell	0.896	0.764	1.051
gaktuell	0.865	0.722	1.036
alter	0.884	0.864	0.905
Geschlecht 2 vs 1	1.319	0.605	2.878
ZENTRUM 2 vs 1	1.327	0.593	2.971
Rauchen 2 vs 1	1.056	0.473	2.357
ATOP_ALL 1 vs 0	3.225	1.442	7.212
ORT 2 vs 0	0.930	0.180	4.800
ORT 1 vs 0	1.295	0.441	3.800

Tabelle 7: Odds Ratio-Schätzer für das Survivalmodell

Alternative Ideen zur Modellwahl

- *Alter* als zeitunabhängige Variable
- *Jahr* als lineare Einflussgröße
- *Alter* als zeitunabhängige Variable und *Jahr* als lineare Einflussgröße

Diese 4 Modelle auch mit den Expositionsscores des jeweiligen Vorjahres.

Auswahl des Modells

- kaum Unterschiede in den Ergebnissen
⇒ **Robustheit**
- Vergleich aller 8 Modelle führt zur Auswahl des **Ausgangsmodells mit den Expositionswerten des jeweils aktuellen Jahres.**

Interpretation des Survivalmodells

- vorhandene Allergie in der Familie steigert Erkrankungsrisiko
- Probanden mit höherem Alter haben niedrigeres Erkrankungsrisiko
Vermutung: Allergien auf Grund von genetischer Vorbelastung
ältere Probanden erkrankten vor der Studie
- leicht negativer Einfluss von hohen Expositionen (nicht signifikant)

Fazit

Negativer Effekt der Schadstoffexpositionen auf Atemwegserkrankungen war nicht nachweisbar!

DANKE

Prof. Leonhard Held

für die Unterstützung bei der Modellierung und Interpretation

PD Dr. Katja Radon

und ihr Team für ausführliche Besprechungen und zahlreiche Erklärungen aus der Welt der Medizin und Epidemiologie