Fall-Kontroll-Studien

Dr. Dirk Hasenclever

IMISE, Leipzig

Hasenclever@IMISE.uni-Leipzig.de

Epidemiologische Grundbegriffe

- **Inzidenz**:
  - Rate des Auftretens von Krankheitsdiagnosen in einer bestimmten Population
  - Oft angegeben in [Fälle/100.000Personen/ Jahr]
  - Quelle: Registerdaten
- **Prävalenz**
  - Rate der zu einem Zeitpunkt von einer Krankheit Betroffenen in der Bevölkerung
  - Quelle: Prävalenz-Querschnittstudien
- **Altersstandardisierung**
  - Krankheitsrisiken sind in der Regel stark altersabhängig
  - Zum Vergleich von Inzidenz/Prävalenz Raten: Standardisierung nötig auf Standard-Alters"pyramiden" ("Europa", "Welt")

Jährliche Inzidenz in Europa von Krebs-Diagnosen und Todesfällen (Krebsregisterdaten)

Internet:: http://www-dep.iarc.fr/eucan/eucan.htm

	Cases	crude	ASR-E	ASR-W	deaths	crude	ASR-E	ASR-W
Oral cavity and pharynx	53556	14.29	12.71	9.28	20178	5.38	4.64	3.31
Oesophagus	24812	6.62	5.38	3.71	22917	6.11	4.85	3.29
Stomach	70798	18.89	14.13	9.35	54919	14.65	10.58	6.81
Colon/Rectum	217526	58.04	44.04	29.36	111781	29.82	21.38	13.63
Liver	31057	8.29	6.41	4.37	34132	9.11	6.81	4.51
Pancreas	41340	11.03	8.35	5.53	45599	12.17	9.02	5.88
Larynx	23304	6.22	5.45	3.92	10326	2.75	2.28	1.59
Lung	196836	52.52	42.16	29.12	183653	49.00	38.27	25.96
Melanoma of skin	38213	10.20	8.89	6.81	9010	2.40	1.94	1.37
Breast	210631	56.20	48.84	35.38	73592	19.63	15.57	10.61
Cervix uteri	22618	6.03	5.35	4.15	10098	2.69	2.17	1.52
Corpus uteri	37411	9.98	8.31	5.81	8998	2.40	1.70	1.08
Ovary etc.	34468	9.20	7.74	5.60	22999	6.14	4.78	3.23
Prostate	144504	38.55	27.77	17.97	56035	14.95	9.73	5.72
Testis	8810	2.35	2.25	2.13	641	0.17	0.15	0.13
Bladder	73132	19.51	14.70	9.78	29773	7.94	5.44	3.35
Kidney etc.	46228	12.33	10.10	7.21	22418	5.98	4.54	3.03
Brain, nervous system	28866	7.70	6.91	5.66	21681	5.78	4.97	3.77
Thyroid	16311	4.35	3.99	3.22	3245	0.87	0.63	0.41
Non-Hodgkin	52440	13.99	11.50	8.46	25906	6.91	5.24	3.55
Hodgkin's disease	8407	2.24	2.13	2.01	2251	0.60	0.49	0.36
Multiple myeloma	21426	5.72	4.36	2.92	15259	4.07	2.93	1.88
Leukaemia	43518	11.61	9.55	7.52	29714	7.93	5.97	4.20
All sites but skin	1580096	421.57	338.83	238.85	929992	248.12	186.54	123.93

Kohorten- versus Fall/Kontrollstudien

- Bei einer **Kohorten-Studie** wird eine feste Studienpopulation im Zeitverlauf beobachtet und **prospektiv** dokumentiert.
- **Kohorten-Studie sind besonders geeignet, um die kausale Wirkung einer Intervention zu prüfen.**
- Kohortenstudien sind unpraktikabel bei seltenen Ereignissen oder Ereignissen, die erst nach langen Wartezeiten auftreten.

- Second best: **Fall/Kontroll Studie**:
  - Finde Fälle mit einer Erkrankung
  - Wähle geeignete Kontrollen ohne die Erkrankung
  - Dokumentiere **retrospektiv** die zu untersuchende Exposition (Rauchen, Asbest etc.)
  - Falls Fälle eine höhere Exposition hatten, Hinweis (aber nicht Beweis!) auf möglichen kausalen Zusammenhang.

Fall-Kontroll Studie: Erhöht spezifische Exposition das Risiko?

Beispiel

Cancer Causes and Control **13**: 609–616, 2002.
© 2002 *Kluwer Academic Publishers. Printed in the Netherlands.*

A case–control study of lung cancer in Casablanca, Morocco

Annie J. Sasco[1,2,*], Ray M. Merrill[1,3], Imane Dari[4], Véronique Benhaïm-Luzon[1], Florence Carriot[1], Cristina I. Cann[1,5] & Mohamed Bartal[4]

[1]*Unit of Epidemiology for Cancer Prevention, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon cédex 8, France;* [2]*Institut National de la Santé et de la Recherche Médicale, Lyon, France;* [3]*Department of Health Science, Brigham Young University, Provo, UT, USA;* [4]*Respiratory Diseases, Ibn Rochd Hospital, Casablanca, Morocco;* [5]*Section of Preventive Medicine and Epidemiology, Boston University School of Medicine, USA*

Fall-Kontroll Studie

Fälle: Lungenkrebs in KH in Casablanca
Kontrollen: 2:1 KH Patienten (andere Diagnose) mit Gleichem Alter, Geschlecht, Wohnort

Standardisiertes Fragebogen Interview.

Demographic variables	Cases (n = 118)	Controls (n = 235)	Chi-square *p*-value
Marital status			
Married	89.0%	93.2%	0.240
Single	3.4%	3.4%	
Divorced or separated	2.5%	0.4%	
Widow(er)	5.1%	3.0%	
Native language			
Arab	97.5%	93.6%	0.122
Berber	2.5%	6.4%	
Literate (can read and write)			
Yes	39.0%	40.4%	0.794
Education			
No formal education	67.8%	64.3%	0.364
Primary school	18.6%	19.2%	
Secondary school	13.6%	14.0%	
University	0.0%	2.6%	
Ethnic background			
Caucasian	91.5%	91.1%	0.851
Black/mixed	8.5%	8.9%	
Place of residence			
Urban	89.8%	90.2%	0.757
Rural	10.2%	9.4%	
Unknown	0.0%	0.4%	
Sex			
Men	96.6%	96.6%	0.994
Women	3.4%	3.4%	
Age at interview			
Mean[a]	59.27	58.83	0.8955
Standard deviation	10.79	10.69	
Median	62	60	
Minimum	35	34	
Maximum	82	82	

[a] Based on *t*-test.

Fall-Kontroll

Quantifikationsproblem
für Expositionsvariable:
(retrospektiv!)
z.B. Rauchen
Pack-years
derzeitiger Status

Resultat:
Multivariat-adjustierte
Schätzer für Odds-Ratios
mittels
logistischer Regression

Table 3. Multivariate matched odds ratios of lung cancer for selected variables, Casablanca, Morocco, January 1996–January 1998

	Matched OR[a]	95% CI
Smoking status		
Never	1.00	Reference
Former light	1.79	0.47–6.79
Former heavy	8.12	2.00–33.06
Current light	18.52	4.10–83.54
Current heavy	26.07	6.58–103.27
Use hashish/kiff and snuff		
None	1.00	Reference
Hashish/kiff	1.93	0.57–6.58
Snuff	1.05	0.28–3.85
Hashish/kiff and snuff	6.67	1.65–26.90
History of chronic bronchitis		
No	1.00	Reference
Yes	4.16	1.76–9.85
Passive smoking[b]		
No	1.00	Reference
Yes	1.36	0.71–2.62
Occupational exposure[c]		
No	1.00	Reference
Yes	1.75	0.84–3.63
Cooking and heat source		
No coal	1.00	Reference
Coal	0.74	0.17–3.14
Lighting source		
No candles	1.00	Reference
Candles	1.44	0.42–5.01
Ventilation of kitchen		
Good	1.00	Reference
Poor	1.22	0.57–2.58

OR: odds ratio; CI: confidence interval.
[a] Adjusted for all other variables listed in table.
[b] Exposure to passive smoking from a spouse, parent, and/or in the workplace.
[c] Exposure to asbestos, iron, nickel, mining activities, fumes, vapors, or other occupational exposures.

Auswertung einer Fall-Kontroll-Studie mittels logistischer Regression I

Erinnerung an logistische Regression in Kohortenstudie :

$$p(D \mid z) = \frac{\exp(\alpha + \beta^T z)}{1 + \exp(\alpha + \beta^T z)},$$

D steht für Erkrankung,

z ist ein Vektor von erklärenden Kovariaten.

mit der Funktion $\text{logit}(p) = \ln\left(\frac{p}{1-p}\right)$ gilt äquivalent :

$$\text{logit}(p(D \mid z)) = \alpha + \beta^T z$$

Auswertung einer Fall-Kontroll-Studie mittels logistischer Regression II

Bei einer Fall - Kontroll - Studie werden Fälle aus
der Population der Erkrankten gezogen
(gesamplet) und entsprechend dazu passende Kontrollen
aus der Population der Nicht - Kranken.
Sei $\pi_1 = pr(S \mid D)$ die bedingte Wahrscheinlichkeit,
dass ein Kranker in die Studie aufgenommen wird.
Sei $\pi_2 = pr(S \mid \neg D)$ die Wahrscheinlichkeit,
dass eine Kontrolle in die Studie aufgenommen wird.

π_1, π_2 sind in der Regel unbekannt.
Daher enthalten die Daten auch keine genaue Information über pr(D | z),
sondern nur über pr(D | S).

Auswertung einer Fall-Kontroll-Studie mittels logistischer Regression III

Nach der Bayes - Regel gilt

$$\mathrm{pr}(D \mid S) = \frac{\mathrm{pr}(S \mid D) * \mathrm{pr}(D)}{\mathrm{pr}(S)} = \frac{\pi_1 * \mathrm{pr}(D)}{\mathrm{pr}(S)}.$$

Ein Individuum ist entweder als Fall oder

als Kontrolle in der Studie : $\mathrm{pr}(S) = \pi_1 * \mathrm{pr}(D) + \pi_2 * (1 - \mathrm{pr}(D))$

Die Erkrankungswahrscheinlichkeit in den Studiendaten

hängt also über die Samplewahrscheinlichkeiten π_1, π_2

mit der unbekannten absoluten

Erkrankungswahrscheinlichkeit zusammen :

$$\mathrm{pr}(D \mid S) = \frac{\pi_1 * \mathrm{pr}(D)}{\pi_1 * \mathrm{pr}(D) + \pi_2 * (1 - \mathrm{pr}(D))}$$

Auswertung einer Fall-Kontroll-Studie mittels logistischer Regression IV

Durch Bildung der Odds erhalten wir :

$$\frac{\text{pr}(D \mid S)}{1 - \text{pr}(D \mid S)} = \frac{\dfrac{\pi_1 * \text{pr}(D)}{\pi_1 * \text{pr}(D) + \pi_2 * (1 - \text{pr}(D))}}{1 - \dfrac{\pi_1 * \text{pr}(D)}{\pi_1 * \text{pr}(D) + \pi_2 * (1 - \text{pr}(D))}} = \frac{\pi_1}{\pi_2} \cdot \frac{pr(D)}{1 - pr(D)}$$

Oder :

$$\text{logit}(\text{pr}(D \mid S)) = \ln\left(\frac{\pi_1}{\pi_2}\right) + \text{logit}(\text{pr}(D))$$

Auswertung einer Fall-Kontroll-Studie mittels logistischer Regression V

Eine logistische Regression mit den Studiendaten liefert also :

$$\text{logit}(\text{pr}(D \mid S,z)) = \ln\left(\frac{\pi_1}{\pi_2}\right) + \text{logit}(\text{pr}(D \mid z)) = \alpha + \beta^T z$$

Die Koeffizienten β lassen sich also als logodds bzgl. der unbekannten Erkrankungswahrscheinlichkeiten in der Gesamtpopulation interpretieren.

!

CAVE: Zufallsbefunde durch Multiples Testen und Subgruppen-Analyse

