

Epidemiologie 1 und Deskriptive Epidemiologie

**Inoffizielles Skript
zu den Vorlesungen des
Studienganges Public Health
der LMU München**

**[Dozenten: E. Grill, R. von Kries, M Wildner,
A. Crispin, A Peters, U. Siebert]**

Stand: 14. Mrz. 2005

Inhaltsverzeichnis

1	Einführung und Hintergründe	3
1.1	Definition „Epidemiologie“:.....	3
1.2	Merkmale der Epidemiologie:	3
1.3	Ziele der Epidemiologie:	3
1.4	Historie.....	5
2	Assoziation und Kausalität	6
2.1	Epidemiologie in der Krankheitsursachenforschung.....	6
2.2	Kriterien für Kausalität	7
2.3	Kausalitätsmodell nach Rothman („Causal Pie Model“, SCN-Modell).....	7
2.4	Kausalität bei Infektionskrankheiten	9
3	Maßzahlen in der Epidemiologie	10
3.1	Maße für Krankheitshäufigkeiten	10
3.2	Epidemiologische Assoziationsmaße (Effektmaße)	12
4	Epidemiologische Studientypen	15
4.1	Deskriptive Studientypen	15
4.2	Analytisch beobachtende Studientypen.....	15
4.3	Analytisch experimentelle Studientypen	17
5	Epidemiologische Datenanalyse	18
5.1	Schritte der Datenanalyse	18
5.2	Rohe, standardisierte, spezifische Maßzahlen	18
5.3	Berechnung der Effektmaße bei Studien.....	18
5.4	Stratifizierte Analyse	21
5.5	Matching	22
5.6	Standardisierung.....	23
6	Fehlerquellen und Störgrößen in epidemiologischen Studien	24
6.1	Klassifikation von Fehlern.....	24
6.2	Selektionsfehler (Selection Bias).....	26
6.3	Informationsfehler (Information Bias).....	27
6.4	Confounding	27
6.5	Effektmodifikation (Interaktion)	29
7	Deskriptive Epidemiologie	30
7.1	Mortalitätsraten	30
7.2	Erkrankungshäufigkeiten in Populationen	33
7.3	Surveillance und Sentinel	36
7.4	Überlebenszeitanalyse	37
8	Anhang	42
8.1	Definitionen.....	42
8.2	Formeln.....	44

1 Einführung und Hintergründe

1.1 Definition „Epidemiologie“:

Epidemiology:

„... the study of the distribution and determinants of disease frequency in human populations“ (MacMahon und Pugh 1970)

Epidemiologie:

Wissenschaftszweig, der sich mit **der Verteilung von übertragbaren und nicht übertragbaren Erkrankungen und deren Determinanten und Folgen in der Bevölkerung** befasst.

1.2 Merkmale der Epidemiologie:

- beobachtende Wissenschaft
- spezifische Fehlerquellen
- Untersuchungen an menschlichen Populationen
- Beobachtung von Krankheiten und Risikofaktoren
- gekennzeichnet durch Interdisziplinarität

1.3 Ziele der Epidemiologie:

- genauer und valider zu beobachten
- Hypothesen zu kausalen Zusammenhängen bilden
- gefährdete Gruppen erkennen
- zukünftige Krankheitsfälle abschätzen
- erkennen „was einen gesund hält“
- Bevölkerungsmedizin vs. Individualmedizin

1.3.1 Epidemiologische Zielebenen / Aufgabenbereiche:

Deskriptive Epidemiologie:

zufällige und systematische Beobachtung bezüglich Populationen,
Subgruppen, Ort und Zeit



Generierung von Hypothesen für einen Zusammenhang zwischen Exposition
und Erkrankung



Analytisch-beobachtende Epidemiologie:

gezielte Überprüfung von Hypothesen zu Zusammenhängen zwischen
Exposition und Erkrankung



Quantifizierung der Effekte von Einflussgrößen



interdisziplinäre Studien zur Beurteilung der Kausalität



Interventionsstrategien



Analytisch-experimentelle Epidemiologie:

gezielte Intervention auf dem Boden der Hypothesen
(meist präventive Interventionen)

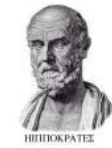


Evaluation der Interventionen

1.4 Historie

„Meilensteine“

500 v.Chr.



Hippokrates: beobachtete verschiedene wichtige und häufige Krankheiten; beschrieb den Einfluss von Klima, Umwelt und Lebensgewohnheiten

1478 - 1553



Girolamo Fracastoro: Verfasser des Gedichtes „Syphilis oder die französische Krankheit“; beschrieb darin den Übertragungsweg der Erkrankung

1620 - 1674



John Graunt: Der englische Kurzwarenhändler führte die erste systematische quantitativ analytische Auswertung von Mortalitätsdaten durch („Bills of mortality“)

<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=200893>

1807 – 1883



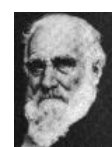
William Farr: Arzt (Assistentin Florence Nightingale), erster Beamter des 1839 gegründeten Büros zur Erstellung von Routinemortalitätsstatistiken für England und Wales. Begründete die erste pragmatische Krankheitssystematik (→ ICD)

1813 – 1858



John Snow: logisch stringente Analyse regionaler Unterschiede in der Wasserversorgung während der Cholera-Epidemie in London 1855

1818 - 1901



Max von Pettenkofer: Choleraepidemie in München 1854, analysierte die hygienischen Bedingungen der Stadt ⇒ Bau einer Kanalisation, Zentralisierung der Schlachthöfe ⇒ Eindämmung der Epidemie

1843 - 1910



Robert Koch: Arzt und Bakteriologe, entdeckte 1882 den Tuberkelbazillus; 1905 Nobelpreis für Medizin und Physiologie; Koch-Henle-Postulate (siehe 2.4.1, Seite 9)

1867-1928



Johannes Fibigers: Experiment 1896/97 bei Diphtheriepatienten (Serum vs. kein Serum, Kopenhagen)

1950



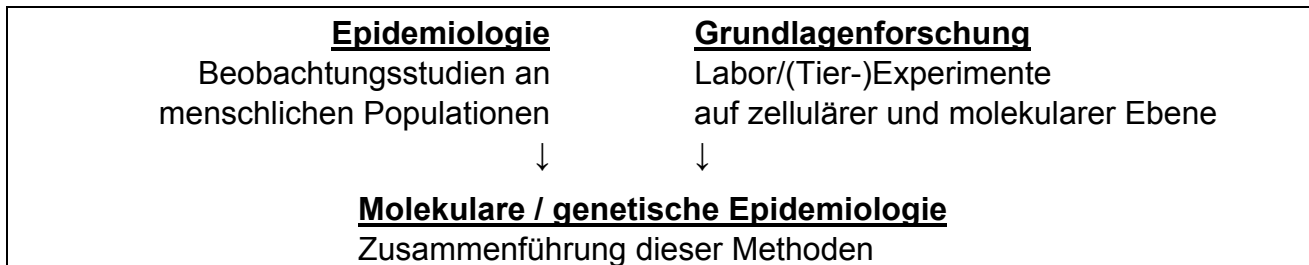
Doll & Hill: Fall-Kontroll-Studie zu Rauchen und Lungenkrebs

1954

Feldstudie Polioimpfung: 1.6 Mio. Schulkinder

2 Assoziation und Kausalität

2.1 Epidemiologie in der Krankheitsursachenforschung



Risikofaktor: wirkt ursächlich (kausaler Zusammenhang)

Risikoindikator: ist mit einem erhöhten Risiko verbunden, ohne dass ein kausaler Zusammenhang vorhanden sein muss

Präventive Maßnahmen:
reduzieren das Erkrankungsrisiko

Beispiele:

Skorbut	Zitrusfrüchte zur Prävention (James Lind, 1747)
Hodenkrebs	Schornsteinfeger als berufliche Risikogruppe (Percival Pott, 1795)
Cholera	Trinkwasserversorgung als Risikoindikator (John Snow, 1855)
Kindbettfieber	Händedesinfektion als präventive Maßnahme (Ignatius Semmelweis, 1874)
Pellagra	einseitige Ernährung als Risikoindikator (Goldberger)
Struma/Kretinismus Jodmangel	Unzureichende Schilddrüsenhormonproduktion als Folge von (1915 Vorschlag zur Jodierung von Speisesalz in der Schweiz)
Minamata-Krankheit	Methylquecksilbervergiftung durch Verzehr vergifteter Fische (WHO, 60er Jahre)
Lungenkrebs	Rauchen als Risikofaktor (Doll und Hill, 1950; Wynder und Graham, 1950)
KHK	Hypertonie, Hyperlipidämie, Rauchen als Risikofaktoren (Framingham-Study, Boston University)
Endometriumkarzinom	Östrogentherapie als Risikofaktor

2.2 Kriterien für Kausalität

Der Nachweis einer statistischen Assoziation zwischen einem vermuteten Risikofaktor und einer Erkrankung belegt noch keinen kausalen Zusammenhang. Daher sollte man zunächst von einer Assoziation zwischen der Exposition (einem Risikofaktor ausgesetzt sein) und der Erkrankung sprechen).

Die Frage nach des Vorhandenseins und der Art des kausalen Zusammenhanges steht jedoch im Mittelpunkt der theoretisch-analytischen Epidemiologie. Um einen einigermaßen sicheren Rückschluss von einer Assoziation auf einen kausalen Zusammenhang zu erreichen wurden Kausalitätskriterien erarbeitet. Die Kausalitätskriterien stehen stets im Kontext aller vorliegenden wissenschaftlichen Erkenntnissen. Trotzdem darf man sie nie kritiklos anwenden.

2.2.1 Bradford-Hill-Kriterien (Sir Austin Bradford Hill, 1963)

1. Stärke der Beziehung	geringe Ausprägung der Assoziation: nur Zufall?
2. Konsistenz der Beziehung	in verschiedenen Studien/Untersuchungen gezeigt?
3. Spezifität des Effektes	klar abgrenzbarer Effekt?
4. Dosis-Wirkungs-Beziehung	höhere Dosis → höhere Wirkung?
5. zeitliche Sequenz	Exposition → Erkrankung
6. Biologische Plausibilität	Ist die Wirkung plausibel?
7. Kohärenz	Widerspruch zum aktuellen Wissenstand?
8. Experimentelle Evidenz	Überprüfbarkeit der Hypothese
9. Analogie mit ähnlichen Kausalzusammenhängen	gibt es Ähnlichkeiten mit anderen Kausalitäten

2.3 Kausalitätsmodell nach Rothman („Causal Pie Model“, SCN-Modell)

S: sufficient cause **C:** component cause **N:** necessary cause

Eine Ursache (cause) ist ein Ereignis, ein Zustand, ein Merkmal oder eine Kombination aus diesen Faktoren

2.3.1 Hinreichende Krankheitsursache (sufficient cause)

Eine Bedingung (einzelne Faktoren oder eine Kombination von mehreren Teil-Faktoren), die unvermeidbar zum Ausbruch einer Krankheit führt

2.3.2 Teilursache (component cause)

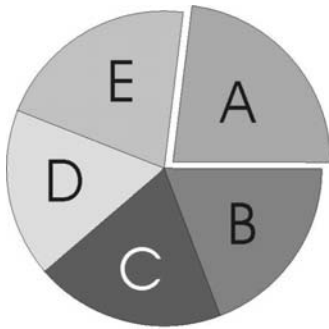
Eine Bedingung, die Bestandteil einer hinreichenden Krankheitsursache ist, für sich alleine aber keine hinreichende Ursache darstellt. Synonyme: beitragende Ursache, Ursachenkomponente

2.3.3 Notwendige Ursache (necessary cause)

Eine Bedingung, die Bestandteil aller hinreichenden Krankheitsursachen ist, d.h. ohne die eine Entstehung der Krankheit nie möglich ist.

Beispiel: Erkrankung mit genau drei hinreichenden Ursachen:

Hinreichende Ursache I



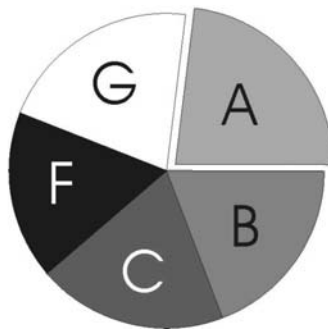
Teilursachen:

A, B, C, D, E, F, G, H, I

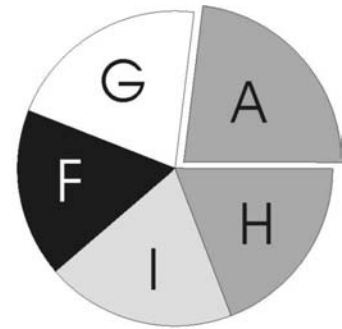
Notwendige Ursache:

A

Hinreichende Ursache II



Hinreichende Ursache III



Beispiel: Tuberkulose

1. Notwendige Ursache: Infektion mit Mykobakterium tuberculosis
2. Teilursachen: schlechte Immunabwehr, anatomisch physiologische Bedingungen
3. Hinreichende Ursache: Zusammentreffen von 1 und 2

2.3.4 Implikationen

Multikausalität

Jede (hinreichende) Ursache besteht aus einer Vielzahl von Teilursachen. Man kann zwischen genetischen (genetic) und äußeren (environmental = non-genetic) Ursachen unterscheiden.

Stärke einer Teilursache

Sie ist abhängig von ihrer Häufigkeit in einer hinreichenden Ursache. Ist sie bei vielen Erkrankungsfällen ein Teil einer hinreichenden Ursache, dann kann sie als stark definiert werden, ist sie hingegen nur bei wenigen Erkrankungen eine Teilursache, dann kann sie als schwach definiert werden. Die Stärke einer Teilursache ist jedoch auch **abhängig von der Häufigkeit** anderer Teilursachen. (Beispiel: Rauchen und Radonexposition. Rauchen ist eine häufige Teilursache bei Lungenkrebsfällen (somit eine starke Teilursache), die Exposition gegenüber Radongas ist dagegen nur bei wenigen Fällen eine Teilursache, also „schwach“. Kann das Rauchen eliminiert werden, so wird die Radongasexposition viel

häufiger eine Teilursache der noch auftretenden Lungenkrebsfälle sein uns somit zu einer „starken“ Teilursache).

Interaktion

Das Zusammentreffen mehrerer Faktoren verursacht Wechselwirkungen zwischen den Teilursachen innerhalb einer hinreichenden Ursache.

Summe der ursächlichen Faktoren

Wird der scheinbare Anteil an einer Erkrankung einzelnen Teilursachen zugerechnet (zuschreibbarer Anteil), dann kann die Summe der zuschreibbaren Anteile (attributables Risiko) über 100% liegen.

Induktionsperiode

Bei jedem Krankheitsfall liegt zwischen dem Auftreten einer Teilursache und dem Beginn der Erkrankung eine gewisse Zeit (Induktionsperiode). Bei epidemiologischen Studien kann der Zeitraum zwischen Exposition und manifester Erkrankung immer nur in Bezug auf einzelne spezifische Teilursache angegeben werden.

2.4 Kausalität bei Infektionskrankheiten

2.4.1 Koch-Henle-Postulate

1. Der krankheitserregende Keim muss regelmäßig von Krankheitsfällen isoliert werden.
2. Er muss in vitro in einer Reinkultur gezüchtet werden können
3. Wenn man eine Reinkultur des Keimes einem Versuchstier (Person) injiziert, muss sich ein typisches Krankheitsbild entwickeln
4. Bei diesem experimentell hervorgerufenen Krankheitsbild muss der Mikroorganismus erneut isoliert werden

2.4.2 Molekular-biologisch nach Stanley Falkow 1988

1. der Phänotyp oder die beobachtete Eigenschaft muss mit pathogenen, nicht aber mit nicht-pathogenen Stämmen des Erregers signifikant korreliert sein
2. gezielte Inaktivierung des/r Gens/Gene für die beobachtete Virulenzdeterminante müssen zu einer signifikanten Verminderung der Pathogenität oder Virulenz führen
3. Die Komplementation mit dem Wildtypen muss zur Wiederherstellung der Pathogenität oder Virulenz führen

3 Maßzahlen in der Epidemiologie

Proportion: [keine Maßeinheit, Range 0 - 1]
$$\frac{\text{Zähler (Teil des Nenners)}}{\text{Nenner}}$$

Rate: Änderung einer Größe im Verhältnis zur Änderung einer anderen Größe, z.B.:
$$\frac{\text{Distanz [km]}}{\text{Zeit [h]}}$$

Ratio: Verhältnis zweier zahlen (z.B. „Krankenhausbetten pro Einwohner“), [keine Maßeinheit]

Odds: „Anteil im Verhältnis zum Komplement“ [keine Maßeinheit, Range 0 - ∞]
$$a / x-a$$

3.1 Maße für Krankheitshäufigkeiten

3.1.1 Prävalenz

Die **Prävalenz** gibt den Anteil kranker Personen zu einem bestimmten Zeitraum (**Punktprävalenz**) oder den Anteil erkrankter Personen innerhalb eines gewissen Zeitraumes in der Bezugsbevölkerung (**Periodenprävalenz**) an. Sie ist eine Proportion und kann als Prozentzahl (zwischen 0 und 100%) oder als Anteil (zwischen 0 und 1) ausgedrückt werden.

$$\text{Prävalenz } P = \frac{\text{Zahl bestehender Fälle}}{\text{Gesamtpopulation}}$$

Wird nur von der Prävalenz gesprochen ist meist die Punktprävalenz gemeint, sie kann als Momentaufnahme in der Bevölkerung aufgefasst werden.

3.1.2 Inzidenz

Die Inzidenz zählt **die Neuerkrankungen innerhalb eines definierten Zeitraumes**, und setzt sie in den Bezug zu den am Anfang dieser Beobachtungsperiode nicht kranken Personen. Sie kann ebenfalls als eine Proportion (Prozentzahl, Anteil) ausgedrückt werden, *bezieht sich aber immer auf einen bestimmten Zeitraum*.

$$\text{Inzidenz } I = \frac{\text{Zahl Neuerkrankungen}}{\text{Gesamtpopulation und Zeit}}$$

3.1.3 Kumulative Inzidenz (Risiko)

Wird für einen definierten Zeitabschnitt und eine bestimmte Anfangsbevölkerung (geschlossene Kohorte) berechnet.

$$CI = \frac{\text{Zahl Neuerkrankungen}}{\text{Bezugsbevölkerung}} \quad [0-100\% \text{ oder } 0-1]$$

3.1.4 Inzidenzdichte (Inzidenzrate)

Bezieht sich nicht mehr auf die Anfangsbevölkerung, sondern auf die im Studienverlauf beobachtete Personenzeit. Diese beginnt für jede Person mit Eintritt in die Studie und läuft entweder a) bis zum Studienende oder b) bis zum Auftreten der Erkrankung oder c) bis zum Tod/Erkrankung aus anderer Ursache (konkurrierendes Ereignis) oder d) bis zum drop out.

$$ID = \frac{\text{Zahl Neuerkrankungen}}{\text{Personenzeit}} \quad [1/10^5 \text{ Personenjahre}]$$

Abschätzen der ID:

$$ID = \frac{a}{(N - a/2 - D/2 - W/2)(\Delta t)}$$

N = Anzahl der Probanden

a = Fälle

D = konkurrierende Ereignisse (z.B. Tod aus anderem Grund)

W = losses to Follow-Up

3.1.5 Zusammenhang zwischen Prävalenz und Inzidenz

Unter bestimmten Voraussetzungen können Inzidenz und Prävalenz ineinander umgerechnet werden:

a) Annäherung:

(Annahme: Zuwanderung zum Prävalenzpool = Abwanderung; D = Dauer der Erkrankung)

$$P \approx ID \cdot D \quad (\text{für niedrige Prävalenzen})$$

b) exakter Zusammenhang

$$P = \frac{ID \cdot D}{1 + ID \cdot D}$$

3.1.6 Zusammenhang zwischen kumulativer Inzidenz und Inzidenzdichte

$$CI = 1 - e^{-\sum ID_i \cdot \Delta t}$$

Ist die ID über die Zeit konstant, dann lässt sich die vereinfachen zu

$$CI = 1 - e^{-ID \cdot \Delta t}$$

Für kleine Werte von $CI < 0.1$ gilt

$$CI \approx ID \cdot \Delta t$$

Δt = Beobachtungsdauer

3.1.7 Odds

$$\text{Odds} = \text{Anteil} / (1 - \text{Anteil})$$

Verhältnis zwischen einem Anteil und dessen Komplement,
z.B. $CI = 0.1$ (10%) \Rightarrow Odds = $1/9$ ($0.111 = 11.1\%$)

3.2 Epidemiologische Assoziationsmaße (Effektmaße)

Effekt ist die Auswirkung einer bestimmten Exposition im Vergleich zur Nicht-Exposition. Man untersucht auf eine Assoziation und deren Stärke zwischen Exposition und Erkrankung (CAVE: Kausalität, Confounding etc.)

Zwei grundlegende Assoziationsmaße:

Verhältnismaße: **RR = Relatives Risiko bzw. Relative Rate**

Differenzmaße **RD = Risikodifferenz bzw. Ratendifferenz**

(Kohortenstudien: Vgl. CI (oder ID) der Erkrankung in Populationen mit und ohne Exposition; Querschnittsstudien: Prävalenz der Erkrankung für Pop. Mit unterschiedlicher Exposition; Fall-Kontroll-Studien: Vergleich der Prävalenz der Exposition zwischen Fällen und Kontrollen)

3.2.1 Verhältnismaße

Relatives Risiko: das Verhältnis der kumulativen Inzidenz zwischen exponierten und nicht-exponierten Personen.

Relative Rate: Vergleich der Inzidenzdichten.

$$RR = \frac{CI_1}{CI_0} \quad \text{bzw.} \quad \frac{ID_1}{ID_0}$$

Odds Ratio: Verhältnis der Erkrankungsodds zwischen Exponierten und Nicht – Exponierten (v.a. bei Fall-Kontroll-Studien)

$$OR = \frac{Odds_1}{Odds_2} \quad (\approx RR)$$

Bei seltenen Erkrankungen ist das Risiko und die Odds zahlenmäßig sehr ähnlich, so dass die RR durch die OR abgeschätzt werden kann (rare-disease-assumption)

3.2.2 Differenzmaße

Risikodifferenz = Attributables Risiko

$$RD = CI_1 - CI_0 = AR$$

Ratendifferenz = Attributable Rate

$$RD = ID_1 - ID_0 = AR$$

Das „überschüssige“ Risiko wird der Exposition zugeschrieben (zuschreibbares Risiko)

Number needed to treat:

$$NNT = 1 / RD$$

3.2.3 Attributable Risiko Proportion unter den Exponierten

= zuschreibbarer Anteil

= ätiologischer Beitrag

= attributable Proportion

= attributable risk percent = attributable fraction = etiologic fraction

ist der Anteil des Krankheitsrisiko, welcher wegfallen würde, bestünde keine Exposition

$$ARP = \frac{I_1 - I_0}{I_1} = 1 - \frac{I_0}{I_1}$$
$$ARP = 1 - \frac{1}{RR} = \frac{RR - 1}{RR}$$

3.2.4 Attributable Risiko Proportion in der Gesamtpopulation

$$\text{PARP} = p_e \cdot \text{ARP} = p_e \cdot (\text{RR} - 1) / \text{RR} \approx p_e \cdot (\text{OR} - 1) / \text{OR}$$

p_e = Anteil der Exponierten in der Gesamtpopulation

3.2.5 Präventive Fraktion für protektive Faktoren

$$\text{PF} = 1 - \text{RR} = 1 - \text{OR}$$

Wobei RR bzw. OR < 1 (protektiver Faktor!)

4 Epidemiologische Studientypen

4.1 Deskriptive Studientypen

Identifizierung und Quantifizierung von wichtigen PH Problemen, Generierung von Hypothesen, Evaluation von Maßnahmen

4.1.1 Kasuistik:

klinische Beobachtungen: Fallbericht, Fallserien \Rightarrow erste Hinweise

4.1.2 Monitoring (räumlich oder zeitlich)

Kontinuierliche Beobachtung und Erfassung bestimmter Merkmale einer bestimmten Bevölkerungsgruppe. Räumliches Monitoring: systematische geographische Erhebung (flächendeckendes Meldesystem). Die erhobenen Daten werden in so genannten Registern verwaltet und analysiert.

Vorraussetzung: *langfristiges und zuverlässiges Meldesystem*

Beispiele: Morbiditätsregister: Herzinfarktregister Augsburg, Krebsregister Saarland
Mortalitätsregister: zeitliches / räumliches soziodemografisches Monitoring (Surveillance, Gesundheitsberichterstattung, medizinische Geografie)

Probleme: Kontinuität der Kodierung; Kontinuität der gesetzlichen Grundlagen; Kontinuität sozioökonomischer Faktoren; Adäquate Daten für Zähler und Nenner; Datenschutz; Kosten

Methoden: Darstellung in Tafelform; Standardisierung (direkt, indirekt)

4.1.3 Ökologische Studie (Korrelationsstudie)

Untersuchen den Zusammenhang zwischen Exposition und Erkrankung auf der Ebene aggregierter Populationen (statt Individuen) und Maße (z.B. per capita Rotweinkonsum – KHK-Mortalität auf Länderebene)

Datengrundlage sind meist routinemäßig erhobene Daten, die oft zu anderen Zwecken erhoben worden waren (z.B. statistische Jahrbücher)

Die Analyse erfolgt bivariat oder mittels multivariater Regressionsverfahren

CAVE: *ökologischer Trugschluss* (z.B. Geburtenrate und Störche)

4.2 Analytisch beobachtende Studientypen

Überprüfung von Hypothesen zur Krankheitsentstehung von Individuen; Identifizierung von Risikofaktoren und Quantifizierung deren Effekte; Entwicklung von Modellen zur Prädiktion von Krankheit und deren Folgen

4.2.1 Querschnittsstudie, Surveys

Simultane Erhebung von Exposition, Erkrankung und Kovariablen

⇒ Assoziation zwischen Prävalenz der Erkrankung und Prävalenz der Exposition

Zensus (Vollerhebung); Stichprobe (Zufallsstichprobe, geschichtete Stichprobe)

Beispiele: Gesundheitssurvey des Bundes; „Gesundheitsbarometer Berlin“

Vorteile: kurze Dauer, multiple Untersuchungen

Nachteile: zeitliche Sequenz unklar, bei seltenen Erkrankungen ineffizient, nur grobe Hinweise auf Kausalbeziehungen

4.2.2 Längsschnittstudien

(= Kohortenstudien, Longitudinalstudien, Follow-Up-Studien)

Gezielte Untersuchung eines Zusammenhanges zwischen Exposition und Erkrankung

prospektiv: Erhebung von Exposition und Kovariablen vor dem Eintreten der Erkrankung

retrospektiv: Erhebung von Exposition und Kovariablen nach dem Eintreten der Erkrankung

Differenzierung nach Art der Studienpopulation:

geschlossene Kohorte (CI oder ID) vs. dynamische Kohorte (immer ID)

Beispiele: Framingham Heart Study
Physicians Health Study
Nurses Health Study

Vorteile: klare zeitliche Sequenz, effizientes Sampling für seltene Exposition möglich, Untersuchung multipler Endpunkte möglich
Bestimmung von absoluten Inzidenzen und Inzidenzdifferenzen möglich

Nachteile: Ineffizient für seltene Erkrankungen, ineffizient für lange Latenzzeiten

4.2.3 Fall-Kontroll-Studien

Vergleich erkrankter und nicht erkrankter Personen bzgl. Exposition und Kovariablen (immer retrospektiv)

Vorgehen: *Krankenhausbezogen:* Definition der Fallserie, Wahl geeigneter Kontrollen
Populationsbezogen: Definition Bezugsbevölkerung, Rekrutierung aller inzidenten Fälle in dieser Bev., Rekrutierung von Kontrollen aus dieser Bev.

Die Schwierigkeit liegt in der Auswahl der Kontrollen!

Vorteile: effizient für seltene Erkrankungen, Untersuchung multipler Risikofaktoren möglich, weniger zeit- und kostenaufwändig

Nachteile. Retrospektive, nicht-blinde Erhebung, Selektion der Kontrollgruppe oft problematisch, ineffizient für seltene Exposition, zeitliche Sequenz nicht immer klar, keine Ermittlung absoluter Inzidenzen möglich, Recall bias

Kontrollen einer Fall-Kontroll-Studie sollten...

- ... bezüglich Expositionsprävalenz repräsentativ für die Bezugsbevölkerung sein, aus der die Fälle stammen
- ... wenn sie erkrankt wären, Teil der Fallgruppe geworden sein
- ... zeitgleich mit den Fällen rekrutiert werden
- ... Kontrollen bleiben, auch wenn sie später erkranken
- ... den gleichen Erhebungsprozeduren unterzogen werden wie die Fälle

zur Durchführung einer Fall-Kontroll-Studie siehe:

Armenian HK, Applications of the Case-Control method, Epidemiologic reviews, 1994, 16

4.3 Analytisch experimentelle Studientypen

4.3.1 Gemeindebezogene Interventionsstudie

Intervention **auf der Ebene von Populationen** (*experimentelle Variante der ökologischen Studie*) meist auf präventiver Ebene.

Ebenen der Prävention:

Primärprävention: ... vor der Krankheitsentstehung (z.B. Impfungen)

Sekundärprävention: ... in der Frühphase der Krankheit (z.B. Brustkrebs-Screening)

Primärprävention: ... zur Vorbeugung der Wiedererkrankung (z.B. Rehabilitation)

4.3.2 Klinische (Interventions-)Studie (Clinical Trial)

Experimentelle Variante der Kohortenstudie auf **individueller Ebene**, möglichst randomisiert und doppelblind (*experimentelle Variante einer Kohortenstudie*).

Rekrutierung: freiwillige Probanden
Informed consent (Aufklärung und Einwilligung)
Studienplan (Ein- und Ausschlusskriterien, Hypothesen etc.)
(möglichst) hohe Probandenzahl
häufig Multicenter-Studien

4.3.3 Quasi-experimentelle Studie

Randomisation by nature, „natürliches Experiment“:

Nutzung der natürlichen Variabilität hinsichtlich Exposition und Outcome zur Analyse (z.B. Radon-Exposition und Lungenkrebs)

5 Epidemiologische Datenanalyse

5.1 Schritte der Datenanalyse

- 1) Studiendesign festlegen, Hypothesen formulieren
- 2) Daten editieren (erfassen, übertragen, Plausibilitätsprüfung)
- 3) Daten zusammenfassen (deskriptive Analyse)
- 4) Daten analysieren

5.2 Rohe, standardisierte, spezifische Maßzahlen

Rohe Analyse: Daten der Studienpopulation insgesamt \Rightarrow rohe Maßzahlen (keine Differenzierung, keine Schichtung)

Standardisierung: direkt oder indirekt

Spezifische Maßzahlen: für Untergruppen

5.3 Berechnung der Effektmaße bei Studien

5.3.1 Kohortenstudie mit PersonENZEITdaten

Rate entspricht der Inzidenzdichte

Risikoschätzer ist die Relative Rate oder die Ratendifferenz

		Erkrankte	Personenzeit
Exposition	+	a	n_1
	-	b	n_0
gesamt		a+b	t

$$\text{Relative Rate } RR = \frac{\text{Rate}_1 (\hat{=} ID_1)}{\text{Rate}_0 (\hat{=} ID_0)} = \frac{\frac{a}{n_1}}{\frac{b}{n_0}} = \frac{a \cdot n_0}{b \cdot n_1}$$

$$\text{Ratendifferenz } RD = \text{Rate}_1 - \text{Rate}_0 (\hat{=} ID_1 - ID_0) = \frac{a}{n_1} - \frac{b}{n_0}$$

5.3.2 Kohortenstudie mit kumulativen Inzidenzen

Risiko entspricht der kumulativen Inzidenz

Risikoschätzer ist das relative Risiko oder die Risikodifferenz

		Erkrankung		gesamt
		+	-	
Exposition	+	a	b	a+b
	-	c	d	c+d
gesamt		a+c	b+d	n

$$\text{Relatives Risiko } RR = \frac{\text{Risiko}_1 (\hat{=} CI_1)}{\text{Risiko}_0 (\hat{=} CI_0)} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}}$$

$$\text{Risikodifferenz } RD = \text{Risiko}_1 - \text{Risiko}_0 (\hat{=} CI_1 - CI_0) = \frac{a}{a+b} - \frac{c}{c+d}$$

5.3.3 Fall-Kontroll-Studie

Risikoschätzer ist nur die Odds Ratio

Keine Inzidenzberechnungen möglich!

		Fälle	Kontrollen	gesamt
		Exposition	+	
-	c		d	
gesamt		a+c	b+d	

$$\text{Odds Ratio } OR = \frac{\text{Odds}_1}{\text{Odds}_0} = \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{a \cdot d}{b \cdot c}$$

rare disease assumption: für seltene Erkrankungen gilt $OR \approx RR$

5.3.4 Querschnittsstudie

Risikoschätzer sind die **relative Prävalenz** oder die **Prävalenzdifferenz**, alternativ die **Odds Ratio**, keine Inzidenzmaßberechnung möglich!

		Erkrankung		gesamt
		+	-	
Exposition	+	a	b	a+b
	-	c	d	c+d
gesamt		a+c	b+d	n

<i>Relative Prävalenz</i>	RP	$=$	$\frac{\text{Prävalenz}_1}{\text{Prävalenz}_0}$	$=$	$\frac{\frac{a}{a+b}}{\frac{c}{c+d}}$
<i>Prävalenzdifferenz</i>	PD	$=$	$\text{Prävalenz}_1 - \text{Prävalenz}_0$	$=$	$\frac{a}{a+b} - \frac{c}{c+d}$
<i>Odds Ratio</i>	OR	$=$	$\frac{\text{Odds}_1}{\text{Odds}_0}$	$=$	$\frac{\frac{a}{c}}{\frac{b}{d}} = \frac{a \cdot d}{b \cdot c}$

für seltene Erkrankungen gilt: $RP \approx OR$

Bei repräsentativen Stichproben kann auch die attributable Risikoproportion ARP und die populationsbezogene ARP (PARP) berechnet werden (siehe 3.2.3, Seite 13)

5.4 Stratifizierte Analyse

Besteht der Verdacht, dass die rohen Effektmaße durch Störgrößen verzerrt sind (Confounding, Effektmodifikation), so kann eine **stratifizierte Analyse** in einzelnen Schichten durchgeführt werden. Hierzu werden stratifizierte Kontingenztafeln erstellt und die stratifizierten (= schichtspezifische) Effektmaße berechnet.

Vorgehen:

Schritt 1: Berechnung der stratumspezifischen Effektmaße

Schritt 2: sind die Effektmaße gleich groß = homogene Assoziation?

NEIN \Rightarrow Effektmodifikation bzw. Interaktion liegt vor

\Rightarrow **Schritt 3:** Stratumspezifische Effektmaße getrennt berichten

JA \Rightarrow keine Effektmodifikation bzw. Interaktion

\Rightarrow **Schritt 3:** gewichteter Durchschnitt der Einzelergebnisse („Pooling“)

5.4.1 Mantel-Haenszel-Verfahren

Dient zur Zusammenfassung schichtspezifischer Effektmaße (Pooling, summary estimator)

<i>Stratum i</i>		Erkrankung		gesamt
		+	-	
Exposition	+	a_i	b_i	a_i+b_i
	-	c_i	d_i	c_i+d_i
gesamt		a_i+c_i	b_i+d_i	$T_i=a_i+b_i+c_i+d_i$

gepoolter Schätzer $RR_{MH} = \frac{\sum \left(\frac{a_i \cdot (c_i + d_i)}{T_i} \right)}{\sum \left(\frac{c_i \cdot (a_i + b_i)}{T_i} \right)}$

bei Fall – Kontroll – Studien $OR_{MH} = \frac{\sum \frac{a_i d_i}{T_i}}{\sum \frac{b_i c_i}{T_i}}$

Um festzustellen, ob die Effekte in den einzelnen Strata signifikant verschieden sind, kann ein statistischer Test durchgeführt werden (Mantel-Haenszel-Teststatistik)

5.5 Matching

Beim Matching werden *zueinander passende Personen* unter Erkrankten und Nicht-Erkrankten (in Fall-Kontroll-Studien) bzw. Exponierten und Nicht-Exponierten (in Kohortenstudien) für die Studienpopulation ausgewählt.

Dadurch sind (in der Regel) keine Rückschlüsse mehr auf die Häufigkeiten/Inzidenzen in der Ausgangspopulation möglich.

Da die Ausprägungen potentieller Störgrößen in den zu vergleichenden Gruppen gleichmäßig verteilt werden, sind validere Aussagen über den spezifischen Einfluss der Exposition möglich.

Hauptanwendung sind Fall-Kontroll-Studien mit kleinen Fallzahlen.

Frequency Matching: lediglich gleiche Verteilung der Matchingvariable in zu vergleichenden Gruppen (z.B. Altersgruppen)

Pair Matching: Zuordnung auf individueller Ebene, Proband \Rightarrow 1 - 4 Matching-Partner

In Kohortenstudien wird so Confounding weitgehend eliminiert, **bei Fall-Kontroll-Studien** ist **zusätzlich eine gematchte Analyse** notwendig

5.5.1 Analyse gematchter Fall-Kontroll-Studien

		Kontrollen		
Fälle		exponiert	nicht-exponiert	Total
exponiert		a	b	a+b
nicht-exponiert		c	d	c+d
Total		a+c	b+d	

Jedes Feld der Kontingenztafel ist mit gematchten Paaren (Fall und dazugehörige Kontrolle) besetzt

$$Odds\ Ratio = \frac{b}{c}$$

In die Berechnung der Odds Ratio gehen nur die divergenten Fälle (b und c) ein!

5.6 Standardisierung (vgl. auch 7.1.3 und 7.1.4, Seite 30)

Ziel: Maßzahlen aus unterschiedlichen Populationen vergleichbar zu machen. Übliche Parameter der Standardisierung sind Alter und Geschlecht.

5.6.1 Direkte Standardisierung

Die *alters und geschlechts-spezifischen Raten* aus verschiedenen Populationen werden auf eine **Standardpopulation** mit definiertem *Altersaufbau und Geschlechtsverhältnis* angewendet.

5.6.2 Indirekte Standardisierung

Auf eine Studienpopulation mit *spezifischem Altersaufbau und Geschlechtsverhältnis* werden die *alters- und geschlechtsspezifischen Raten* einer **Vergleichsbevölkerung** angewendet.

Die indirekte Standardisierung erlaubt den Vergleich zweier Bevölkerungen ohne Verwendung einer Standardbevölkerung.

6 Fehlerquellen und Störgrößen in epidemiologischen Studien

6.1 Klassifikation von Fehlern

Finden sich in einer epidemiologischen Studie Assoziationen so können diese durch Fehler beeinflusst oder gar vorgetäuscht werden. Es gibt zwei grundlegende Klassen von Fehlern:

1. Zufällige Fehler
2. Systematische Fehler

6.1.1 Zufällige Fehler

Da die Epidemiologie mit Stichproben arbeitet, sind die Effektschätzer mit zufälligen Fehlern behaftet. Eine Quantifizierung der Fehlermöglichkeiten ist möglich durch:

1. Statistische Tests („testen“)
2. Konfidenzintervalle („schätzen“)

Dualität von Schätzen und Testen

Schätzen: unbekannte Parameter der GG anhand einer gezogenen Stichprobe möglichst genau festlegen \Rightarrow Konfidenzintervall

Testen: Hypothese aufstellen und anhand der Daten entscheiden, ob die Hypothese angenommen werden kann oder abgelehnt werden muss \Rightarrow Entscheidung richtig oder falsch (\Rightarrow p-Wert)

Definition des p-Wertes:

Ein p-Wert ist die Wahrscheinlichkeit, das (vorliegende) Studienergebnis oder ein noch extremeres Ergebnis zu erhalten, wenn die Hypothese H_0 wahr ist.

Definition Konfidenzintervall:

Das 95%-Konfidenzintervall bedeckt den Parameterwert mit einer Wahrscheinlichkeit von 0.95 (analog 99%-KI-Intervall etc.)

6.1.2 Systematische Fehler

Systematische Fehler führen zu einer **Verzerrung (Bias)** der Effektschätzer (RR, OR etc.) Sie sind daher in der Epidemiologie von überragender Bedeutung.

Zufällige Fehler	erniedrigen die	Präzision
Systematische Fehler	erniedrigen die	Validität

Die Validität hat Priorität vor der Präzision.

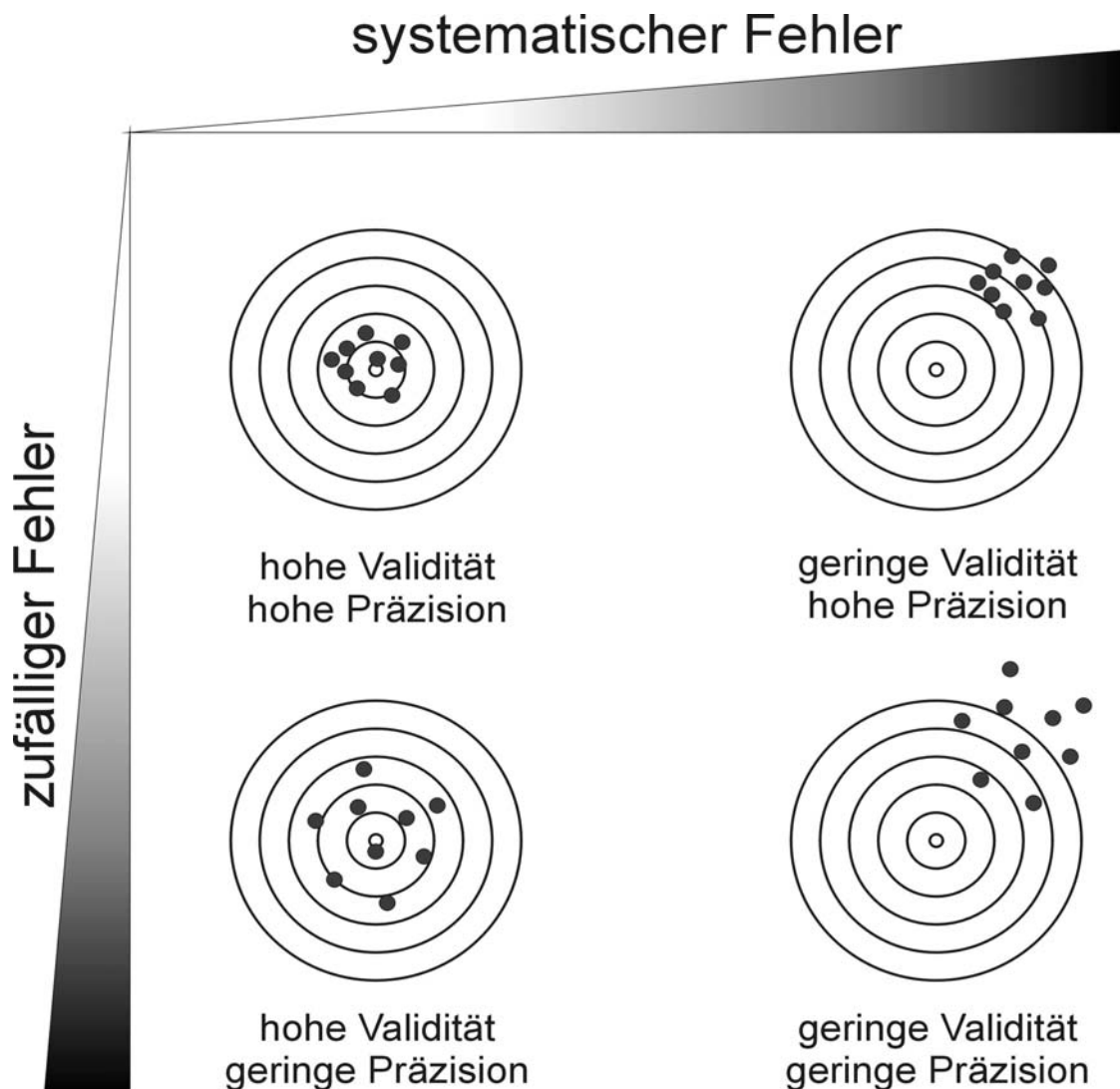


Abbildung 1: Validität und Präzision in Abhängigkeit von zufälligem und systematischem Fehler

6.1.3 Formen der systematischen Fehler

Entstehung des Fehlers	Bezeichnung	Beispiel
Bei der Rekrutierung und beim Follow-up der Studienteilnehmer	Selektionsfehler (Selection Bias)	Nichterreichen von Berufstätigen bei Telefonumfragen
Bei der Erhebung der Merkmale (Exposition, Kovariablen, Erkrankungsstatus)	Informationsfehler (Information Bias)	Fehlklassifikation / Fehldiagnosen, Messfehler
Mangelnde Berücksichtigung von Störgrößen	Confounding	Scheinassoziation zwischen Alkohol und Bronchialkarzinom, über Rauchen erklärbar

Besonderheiten des Confounding:

Die Zuordnung des Confounding ist nicht einheitlich. Entweder man ordnet Confounding den systematischen Fehlern (Bias) zu, da es zu einer verzerrten Schätzung des (wahren) Effektmaßes führt.

Alternativ zählt man nur den Selektions- und den Informationsfehler zu den systematischen Fehlern und betrachtet das Confounding als eigenes Phänomen, da Selektions- und den Informationsfehler bei der Studienplanung/Datenerhebung zumindest teilweise vermieden werden können, während Confounding ein in der Natur der Grundgesamtheit verankertes Phänomen darstellt und nur durch Randomisierung vermieden bzw. in der Analyse korrigiert werden kann.

6.2 Selektionsfehler (Selection Bias)

Hauptquellen:

Querschnittsstudie: verfehlte Kontaktaufnahme, Teilnahmeverweigerung, nicht verwertbare Teilnahmebögen, etc.

Fall-Kontroll-Studie: unvollständige Rekrutierung der Fälle, nicht repräsentative Kontrollen für die Bezugspopulation

Kohortenstudie: siehe Querschnittsstudie, Follow-up: Wegzug, Sterbefälle, Verweigerung etc. (drop-out)

Gegenmittel:

Alle: Wiederholte Kontaktaufnahme
Multiple Zugangsverfahren (Telefon, Post, Register, etc.)

Beispiele:

Healthy-Worker-Effekt, Ill-worker-effect, Middle-Class bias

6.3 Informationsfehler (Information Bias)

Hauptquellen:

Erhebung von Exposition und Kovariablen:

- Gedächtnisprobleme (auch: „Recall Bias“ bei Fall-Kontroll-Studie)
- Bewusste Fehlangaben (z.B. bei sozial geächteten Gewohnheiten)
- Falsches „Zeitfenster“, falsche Messtechnik, Rundungsfehler
- Ziffernpräferenz, Übertragungsfehler
- „Zentrale Tendenz“ bei Fragebögen (extreme Antworten bei vorgegebenen Antwortmöglichkeiten werden seltener verwendet)

Erhebung des Erkrankungsstatus:

- Follow-up-Mechanismen unvollständig (z.B. Krebsregister)
- „stumme“ Erkrankungen (geringe Sensitivität der Krankheitserkennung)
- Fehldiagnosen durch Erwartungshaltung
- Unspezifische Symptome (geringe Spezifität der Krankheitserkennung)

Gegenmittel:

- Standardisierte Erhebung der Daten
- Schulung der Mitarbeiter
- Multiple unabhängige Erhebungen
- Validierungsstudien
- Pretests
- Plausibilitätskontrollen

Beispiele:

- Recall bias

6.4 Confounding

6.4.1 Definitionen:

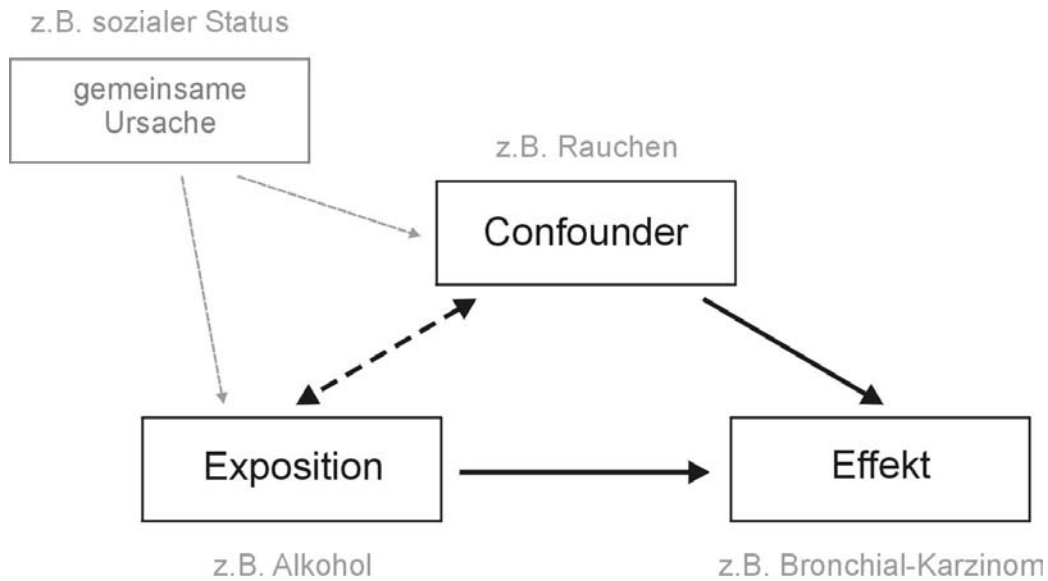
Klassische Definition:

Ein Confounder ist eine dritte Variable, die mit der interessierenden Exposition assoziiert auftritt und einen unabhängigen Einfluss auf den Effekt ausübt, jedoch keine intermediäre Variable ist.

Change-in-estimate Definition (Robins):

Ein Confounder ist eine Variable, deren Berücksichtigung (Kontrolle) zu einem anderen Effektschätzer führt.

Abbildung 2: „Confounding-Dreieck“ (modifiziert)



Es besteht eine vermeintliche Assoziation zwischen der Exposition und des Effektes (z.B. Krankheit) aufgrund einer Störgröße (Confounder), die sowohl mit der Exposition als auch mit dem Effekt (Erkrankung) assoziiert ist. Confounder und Exposition können eine gemeinsame Ursache haben (dies ist jedoch nicht zwingend der Fall).

Der Confounder kann die Assoziation in jeder Richtung beeinflussen.

Beispiele für Confounder:

- | | |
|---|--------------------------|
| Mortalitätsunterschiede Alaska – Alabama: | Alter |
| Alkohol – Lungenkrebs: | Rauchen |
| Rauchen – Osteoporose: | Geschlecht |
| Stadtteil – Mortalität: | Sozioökonomischer Status |

Intermediärvariable:

Eine Zwischenstufe einer Kausalkette ist *kein Confounder* sondern eine *Intermediärvariable*.

Abbildung 3:

Intermediärvariable: Zwischenstufe einer Kausalkette Exposition → Intermediärvariable → Effekt



6.4.2 Gegenmittel gegen Confounding

1) auf der Ebene der Studienplanung

- Restriktion (bezüglich der Einschlusskriterien)
- Matching (\Rightarrow bezüglich Kovariablen übereinstimmende Gruppen)
bei Fall-Kontroll-Studien gematchte Analyse nötig!
- Randomisierung (\Rightarrow Kontrolle auch für unbekannte Störgrößen)

2) auf der Ebene der Analyse

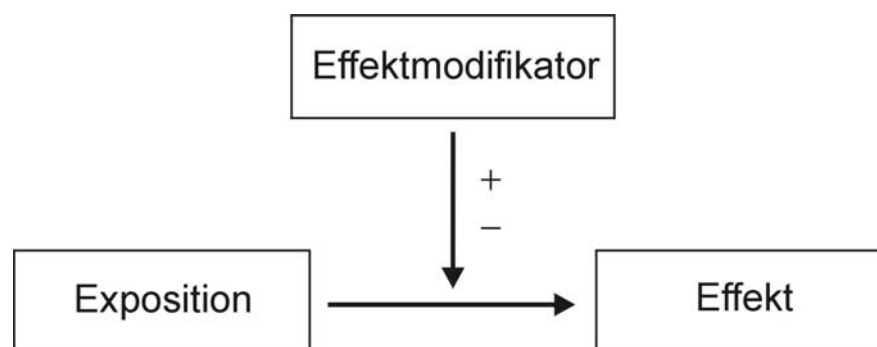
- Stratifizierte Analyse (Schichtung nach Ausprägung der Kovariablen)
- Anwendung multivariater statistischer Modelle
(z.B. multiple Regressionsanalysen, Cox-Regression, log-odds)

Vorraussetzung ist die akkurate Messung der relevanten Störgrößen (außer bei der Randomisierung)

6.5 Effektmodifikation (Interaktion)

Existieren in verschiedenen Strata (Schichten) einer Variablen unterschiedliche Effektschätzer, so spricht man von **Effektmodifikation** bzw. **Interaktion**. Die Schichtvariable wird als Effektmodifikator (effect modifier) bezeichnet. Reine Effektmodifikation führt nicht zu einer Verzerrung des Effektmaßes und gehört damit nicht zu den Fehlern in epidemiologischen Studien.

Abbildung 4: Effektmodifikation



Beispiel:

Rauchen, Lungenkrebs und Radonexposition

ARP(Radon) für Raucher = 4 – 7 %

ARP(Radon) für Nichtraucher = 14 – 22 %

\Rightarrow Wechselwirkung (Interaktion) zwischen Tabakrauchen und Radon

\Rightarrow Maßzahlen sind gesondert ($\hat{=}$ stratifiziert) zu berichten

7 Deskriptive Epidemiologie

7.1 Mortalitätsraten

Bestimmung der Zahl der Sterbefälle in einer definierten Population über einen definierten Zeitraum

- Sterbefälle/Bevölkerung/100.000
- Relevant in der Demografie: Sterberate > Geburtenrate \Rightarrow Bevölkerung schrumpft
- Gesundheitspolitisch wichtig: Hohe regional Mortalitätsraten können besondere lokale Gesundheitsrisiken anzeigen

7.1.1 Altersspezifische Mortalität (Sterblichkeit)

Gompertzches Gesetz:

Beziehung von Sterblichkeit und Alter, die gekennzeichnet ist durch:

- höhere Sterblichkeit im ersten Lebensjahr
- einem anschließenden Abfall
- einem logarithmischen Wiederanstieg ab der Altersgruppe 35 bis 40

7.1.2 Stratumspezifische Mortalitätsrate

- Stratum = Schicht (z.B. Altersgruppe 5 – 9 Jahre)
- MR_i = Mortalitätsrate des Stratum i
- Die Summe *der stratumspezifischen Mortalitätsraten* ist die *Mortalitätsrate*

$$\sum MR_i = MR$$

7.1.3 Direkte Standardisierung der Mortalitätsraten

Prinzip:

Die stratumspezifischen MR der Indexpopulation werden auf die Besetzung des Stratum in der Standardpopulation bezogen

Wie viele Fälle/Tote wären es in der Indexpopulation gewesen, wenn die Indexpopulation die gleiche Altersstruktur wie die Standardpopulation hätte“)

Notwendig:

Standardpopulation: Altersstruktur
Indexpopulation: Sterberaten nach Alter

nicht unbedingt notwendig:

(Sterberaten nach Alter)
(Alterstruktur)

Formeln:

$$\text{Standardisierte MR} = \frac{\sum N_i \cdot p_i}{\sum N_i} \cdot 10^5$$

N_i = Anzahl der Individuen in der i ten Altersgruppe der Standardpopulation

p_i = Altersspezifische Sterberate der Indexpopulation

$$\text{Standard error SE} = \sqrt{\frac{\sum \frac{N_i^2 \cdot p_i}{n_i}}{(\sum N_i)^2}}$$

N_i = Anzahl der Individuen in der i ten Altersgruppe der Standardpopulation

p_i = Altersspezifische Sterberate der Indexpopulation

n_i = Anzahl der Individuen in der i ten Altersgruppe der Indexpopulation

7.1.3.1 CMF: Comparative Mortality Figure

Gibt an, um wie viel die standardisierte MR in der Indexpopulation höher ist als die MR in der Standardpopulation bzw. die standardisierte MR der verglichenen Indexpopulationen

$$CMF = \frac{MRa_{st}}{MRsp} \quad \text{bzw.} \quad CMF = \frac{MRa_{st}}{MRb_{st}}$$

MRa_{st} : standardisierte MR der Indexpopulation a

$MRsp$: MR der Standardpopulation

MRb_{st} : standardisierte MR der Indexpopulation b

7.1.4 Indirekte Standardisierung der Mortalitätsraten

Prinzip:

Die stratumspezifischen MR der Standardbevölkerung werden mit der Besetzung des Stratums in der Indexpopulation gewichtet

„Wie viele Fälle/Tote würde ich in der Indexpopulation erwarten, wenn die Sterblichkeit in der Indexpopulation so hoch wäre wie in der Standardpopulation“

Notwendig:

Standardpopulation: Sterberaten nach Alter

Indexpopulation: Altersstruktur

nicht unbedingt notwendig:

(Altersstruktur)

(Sterberaten nach Alter)

Formeln:

$$\text{erwartete Fälle} = \sum P_i \cdot n_i$$

n_i = Anzahl der Individuen in der i -ten Altersgruppe der Indexpopulation

P_i = Altersspezifische Sterberate der Standardpopulation

7.1.5 Standardisierte Mortalitäts Ratio (SMR) Standardisierte Inzidenz Ratio (SIR)

$$SMR = \frac{\text{Zahl der beobachteten Fälle in der Indexpopulation}}{\text{Zahl der erwarteten Fälle in der Indexpopulation}}$$

$$SMR = \frac{\sum r_i}{\sum n_i \cdot P_i}$$

$$SE = \frac{SMR}{\sqrt{\sum r_i}} \quad \text{Standardfehler}$$

r_i = Anzahl der Sterbefälle in der Indexpopulation

n_i = Anzahl der Individuen in der i -ten Altersgruppe der Indexpopulation

P_i = Altersspezifische Sterberate der Standardpopulation

Entsprechend lässt sich auch die **Standardisierte Inzidenz Ratio** für Erkrankungen berechnen:

$$SIR = \frac{\sum r_i}{\sum n_i \cdot P_i}$$

Beispiel SIR:

Zöliakie (Glutenunverträglichkeit):

- Seltene Erkrankung (ca. 1:1300 bis 1:5000)
- Klinisches Krankheitsbild:
Kindesalter: Gewichtsverlust, Durchfälle, Hypotonie, Missmutigkeit
Erwachsene: Minderwuchs, Entwicklungsretardierung, neurologische Symptome, Vitaminmangelerkrankungen, Hautsymptome
- Erhöhtes Lymphomrisiko (?)

⇒ zur Klärung des vermuteten Lymphomrisikos: SIR berechnen

Daten zur Erkrankung aus Registerstudien (Schweden, Askling et al., Gastroenterology 123: 1428): in Schweden einheitliche Registriernummer für ärztliche Behandlungen, verlässliche ICD-Kodierung, Krebsregister

⇒ Identifikation der Fälle (Krankenhausstatistik)

⇒ Follow-up bis Diagnose Lymphom, Tod andere Ursache, etc.

⇒ Berechnung der SIR möglich

7.2 Erkrankungshäufigkeiten in Populationen

Probleme:

- Wie können Fälle erfasst werden?
- **Diagnostische Bias:** Fall wird nicht diagnostiziert
- **Underreporting:** Diagnostizierter Fall wird nicht gemeldet

7.2.1 Erfassung von Krankheitsdaten/fällen

Passive Meldesysteme:

- Meldepflicht (z.B. gesetzliche Vorgaben bei bestimmten Infektionen)
- Freiwillige Meldungen

Aktive Erfassungssysteme:

- Nachfragen
- z.B. durch Fragebögen

Beispiel ESPED-Modell:

Hatten Sie einen Fall?

Rückmeldung auch wenn kein Fall gemeldet wurde
neben Meldekarten auch Fragebögen

⇒ fast vollständiger Rücklauf der Meldekarten und Fragebögen

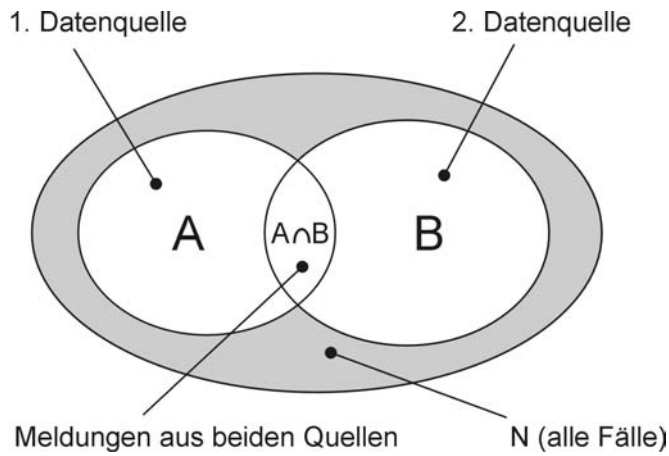
95% Rücklauf der Meldekarten

95% Rücklauf der Fragebögen

⇒ errechnete Vollständigkeit $0.95 * 0.95 = 0.90$

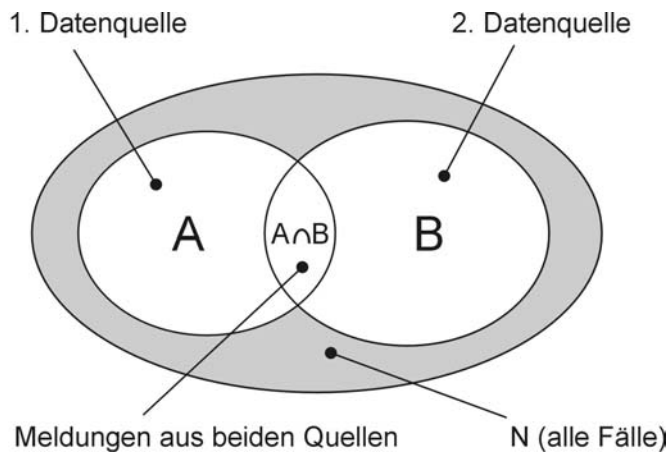
7.2.1.1 Vollständigkeit der Erfassung von Daten: Capture-Recapture-Prinzip

Traditioneller Ansatz:



⇒ Problem der Schätzung des tatsächlichen N

Capture-Recapture-Prinzip:



$$\frac{A}{N} \cdot \frac{B}{N} = \frac{(A \cap B)}{N}$$

$$\Rightarrow A \cdot B = (A \cap B) \cdot N$$

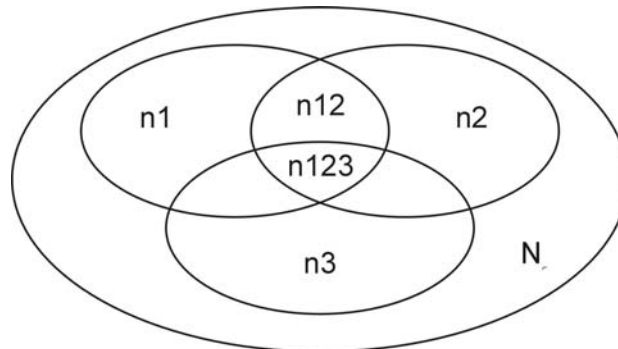
$$N = \frac{A \cdot B}{(A \cap B)}$$

Vorraussetzungen für das Capture-Recapture-Prinzip:

- Geschlossene Population
- Perfektes Linking (Linking zu groß \Rightarrow N zu klein; L zu klein \Rightarrow N zu groß)
- Unabhängige Erfassung
 $\Rightarrow A/N \cdot B/N = (A \cap B)/N$ (unabhängige Wahrscheinlichkeiten)
- Keine Heterogenitäten (\Rightarrow Erfassung der gleichen Population)

Abschätzung der Abhängigkeiten der Erfassung:

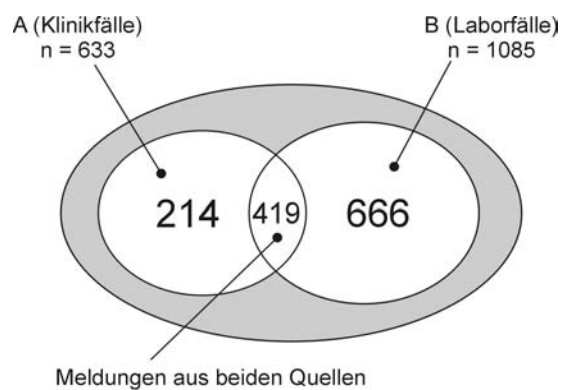
Dritte Datenquelle:



$$\frac{n_{12}}{N_c} \cdot \frac{n_3}{N_c} = \frac{n_{123}}{N_c}$$

Beispiel:

Pneumokokken-Infektionen (vor Einführung einer Impfung)



erfasste Fälle gesamt = traditionelle Schätzung von N:

$$N = A + B - (A \cap B) \quad N = 633 + 1085 - 419 = 1299$$

Capture-Recapture:

$$N = \frac{A \cdot B}{(A \cap B)} \quad N = \frac{633 \cdot 1085}{419} = 1639$$

erfasste Fälle: $1299/1639 = 79\%$

Zusammenfassung:

- Mit dem *Capture-Recapture-Prinzip* lässt sich (wenn alle Bedingungen ideal sind) ein *Reporting Bias* ausgleichen
⇒ eine korrektere Schätzung der diagnostizierten Fälle ist möglich
- Der Diagnostische Bias ist dadurch nicht auszugleichen

7.3 Surveillance und Sentinel

Surveillance:

Systematische fortwährende Sammlung und Analyse von Daten und (kurzfristiges) Feedback an diejenigen, die daraus Rückschlüsse ziehen und darauf reagieren können (sollten): „Daten für Taten“

Beispiele:

Herzinfarktodesfälle pro Jahr; HIV-Neuinfektionen/Prävalenz bei Frau in gebärfähigen Alter – pro Jahr; Masernerkrankungsrate pro Jahr (Erkennung von Zyklen); Influenza pro Woche (wann beginnt die Saison?, Behandlungsempfehlungen)

Sentinel:

- Bestimmte Krankheit als Indikator (z.B. Schädel-Hirn-Verletzungen als Indikator für Verletzungen bei Verkehrsunfällen)
- Sentinel Praxis: Berichtspraxis, die bestimmte Erkrankungen melden (auf freiwilliger Basis), Frage: sind die Praxen repräsentativ?
- Sentinel surveillance: kontinuierliche Erfassung der Meldungen der Sentinel Praxen
- Welche Bezugsgrößen nimmt man? (Fallzahlen pro Zeitraum? Fallzahlen pro Zeitraum und Praxis? Fallzahlen pro Zeitraum und Behandlungen?)
- Erkrankungen, bei denen Patienten nicht zum Arzt gehen werden nicht erfasst

Sentinel surveillance - Beispiele:

Laufende Programme: Masern, Influenza (www.influenza.rki.de/agi), Varizellen (in Vorbereitung)

Arbeitsgemeinschaft Influenza AGI:

Ziele:

wöchentliche Erfassung von Erkrankungen, Arbeitsunfähigkeiten, Krankenhauseinweisungen; Erkennen neuer Varianten der Erreger; Passung des Impfstoffes auf das relevante Erregerspektrum des Jahres

Instrumente:

Praxis Sentinel, virologische Analysen von Rachenabstrichen im NRZ (nationales Referenzzentrum), Meldungen nach dem IFSG (Infektionsschutzgesetz)

Sentinel Praxen: 656 Allgemeinmediziner, Pädiater und Internisten, wöchentliche Meldungen, Falldefinition (akute respiratorische Infektionen), Bezugsgröße: Praxiskontakte pro Woche; 1/3 der Praxen kann Rachenabstriche nehmen

7.4 Überlebenszeitanalyse

Beispiele für Fragestellungen:

Wie lange überleben Krebspatienten ab Diagnosestellung?

Wie lange ist das erkrankungsfreie Überleben bis zu Progression der Erkrankung?

Wie lange überleben Patienten nach Herztransplantation?

Wie lange ist die Dauer bis zum Verschluss eines Koronarstent?

Wie lange wartet ein Dialysepatient bis zur Nierentransplantation?

etc.

Warum sind spezielle Methoden für Survival-Analysen notwendig?

Auf den ersten Blick ist die Überlebenszeit eine gewöhnliche quantitative Variable (auf die deskriptiv-statistische Verfahren angewendet werden können: Mittelwert, Median, Standardabweichung ...).

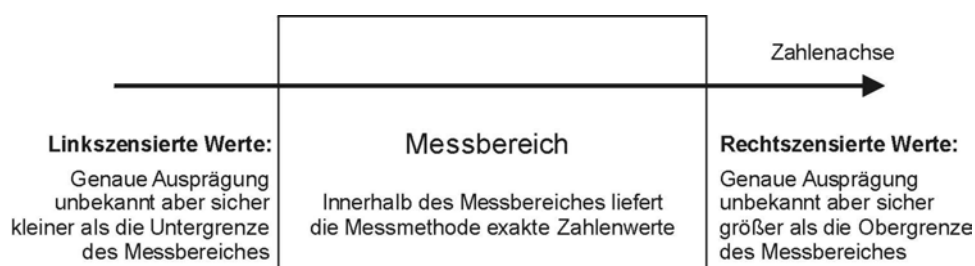
Beispiel: London Times 30.3.1990: Verkehrspiloten sterben jung: 60% der Verstorbenen waren zum Zeitpunkt des Todes jünger als 65 Jahre.

Wo ist der Denkfehler?

Die bereits Verstorbenen sind nicht repräsentativ für die Gesamtpopulation. Werden nur die Überlebenszeiten der Verstorbenen betrachtet, so resultiert eine verzerrte (in der Regel zu pessimistische) Schätzung. Die einfache Berechnung von Mittelwert und Streuung erbringt nur dann valide Ergebnisse, wenn die Überlebenszeiten *aller* Mitglieder der untersuchten Population bekannt sind. Dies ist jedoch so gut wie nie der Fall.

Aussagen zur Lebenserwartung betreffen Lebende, die erst in der Zukunft versterben werden. Um Aussagen zur Lebenserwartung machen zu können kann man nicht warten bis alle Populationsmitglieder verstorben sind (auch in Kohortenstudien (Follow-Up-Studien) bei ungünstigen Erkrankungen (z.B. Tumore) überleben viele Patienten das Studienende. Zudem gibt es Teilnehmer, die während der Studie ausscheiden (Losses to Follow-up): In beiden Fällen ist die *Überlebenszeit unbekannt*. Bekannt ist jedoch wie lange die Patienten beobachtet wurden \Rightarrow die Überlebenszeit ist mindestens so lang wie die Beobachtungszeit. Im Fachjargon heißt dies: Die Lebenszeiten sind *zensiert*.

Abbildung 5: Zensierung von Daten:



Überlebenszeiten können rechtszensiert sein!

7.4.1 Perioden- und Kohortensterbetafeln

Kohortensterbetafeln:

- Beschreibung des **Absterbens realer Kohorten**
- Datengrundlage: **Längsschnittdaten**, die an echten Patientenstichproben gewonnen wurden
- Einsatzgebiet: **Klinische Epidemiologie** (z.B. Studien, Tumorregister)

Periodensterbetafeln:

- Beschreibung des **hypothetischen Absterbens einer hypothetischen Kohorte** von 100.000 Neugeborenen
- Datengrundlage: **aktuelle alterspezifische Mortalitätsraten**, also **Querschnittsdaten**
- **Hilfsannahme**: konstante, zukünftig **unveränderte Mortalitätsrisiken**
- Einsatzgebiet: Abschätzung der **Lebenserwartung** in der **Bevölkerungsstatistik**

Probleme der Periodensterbetafeln:

- In Deutschland *sinken die alterspezifischen Mortalitätsraten* seit Jahren stetig \Rightarrow die Periodensterbetafeln haben die *Lebenserwartung* daher bislang immer *unterschätzt*
- In anderen Ländern (z.B. Botswana) *steigen die Mortalitätsraten* seit einigen Jahren an (u.a. wegen HIV) \Rightarrow die Periodensterbetafeln haben die *Lebenserwartung* *überschätzt*

7.4.2 Konstruktion von Periodensterbetafeln

l_x	Anzahl der Personen, die das Alter x erreichen ($l_0 = 100.000$)
m_x	Altersspezifische Mortalitätsrate pro 1000 Personenjahre
q_x	Sterberisiko im Intervall $[x, x+t]$
p_x	Wahrscheinlichkeit, das Intervall $[x, x+t]$ zu überleben
L_x	Gelebte Jahre im Intervall $[x, x+t]$
T_x	Gelebte Jahre nach Erreichen von x
e_x	Lebenserwartung im Alter x

Die **alterspezifische Mortalitätsrate** m_x ist die Inzidenzdichte von Todesfällen in der Altersgruppe $[x, x+t]$

Die **Sterbewahrscheinlichkeit** q_x ist die kumulative Inzidenz von Todesfällen im Zeitraum $[x, x+t]$

Sterbewahrscheinlichkeit q_x

Die Formel zur Umrechnung von Inzidenzdichten in kumulative Inzidenzen lautet:

$$CI = 1 - e^{-ID_i \cdot \Delta t}$$

$$\Rightarrow q_x = 1 - e^{-m_x \cdot \Delta t}$$

Überlebenswahrscheinlichkeit p_x

$$p_x = 1 - q_x$$

Überlebende l_x im Alter x

$$l_0 = 100.000$$

$$l_{0+t} = l_0 \cdot p_0$$

$$l_{x+t} = l_x \cdot p_x$$

Gesamtzahl der gelebten Jahre L_x im Intervall $[x, x+t]$

Personen, die das Intervall $[x, x+t]$ überleben, bringen jeweils t gelebte Jahre ein:

$$l_{x+t} \cdot t$$

Personen, die im Intervall $[x, x+t]$ verstorben sind, bringen im Mittel $t/2$ gelebte Jahre ein:

$$(l_x - l_{x+t}) \cdot \frac{t}{2}$$

$$\Rightarrow L_x = l_{x+t} \cdot t + (l_x - l_{x+t}) \cdot \frac{t}{2}$$

$$L_x = (l_x + l_{x+t}) \cdot \frac{t}{2}$$

Gelebte Jahre T_x nach Erreichen des Alters x

Summation aller gelebten Jahre aus allen Altersgruppen größer oder gleich x

$$T_x = \sum_{i \geq x} L_i$$

Lebenserwartung im Alter x

Durchschnittliche Zahl der gelebten Jahre pro Person im Alter x

$$e_x = \frac{T_x}{l_x}$$

7.4.3 Konstruktion von Kohortensterbetafeln

Cutler-Ederer-Verfahren

= Sterbetafelmethode = Aktuarsmethode

- Die individuelle Beobachtungszeit beginnt mit dem Eintritt in die Studie
- Die Beobachtungszeit wird in fixe Intervalle unterteilt
- Die Intervalle werden mit der laufenden Nummer x durchnummeriert
- Am Anfang jedes Intervalls x_i wird die **Zahl l_x der Überlebenden** bestimmt
- Am Ende jedes Intervalls x wird die Zahl der **Todesfälle d_x** und der **Zensierungen a_x ermittelt** (Drop-out oder Ende der Beobachtungszeit)

l_x	Lebende Probanden zu Beginn des Intervalls x
d_x	Sterbefälle im Intervall x
a_x	Zensierte Fälle im Intervall x
L_x	Beobachtete Personen unter Risiko im Intervall x
q_x	Sterberisiko im Intervall x
p_x	Überlebenswahrscheinlichkeit im Intervall x
S_x	Kumulative Überlebenswahrscheinlichkeit bis zum Ende des Intervalls x

Sterbewahrscheinlichkeit q_x

Ohne Zensierungen ($a_x = 0$) könnte man q_x einfach berechnen, indem man die d_x Verstorbenen auf die l_x Lebenden zu Beginn des Intervalls bezieht

$$q_x = \frac{d_x}{l_x}$$

Bei **Zensierungen** ist dieser Ansatz jedoch **zu optimistisch**, da nicht alle l_x Personen während des gesamten Intervalls unter Risiko stehen, einen beobachtbaren Tod zu erleiden!

Zahl der beobachteten Personen unter Risiko L_x

Hilfsannahme: Zensierungen erfolgen im Mittel in der Mitte des Intervalls. Daher zählen Zensierungen nur halb

$$L_x = l_x - \left(\frac{a_x}{2} \right)$$

Sterbewahrscheinlichkeit q_x und Überlebenswahrscheinlichkeit p_x im Intervall

$$q_x = \frac{d_x}{L_x} \qquad p_x = 1 - q_x$$

Survival: Kumulative Überlebenswahrscheinlichkeit

Zu Beginn der Beobachtung ist die kumulative Überlebenswahrscheinlichkeit 100%:

$$S_0 = 1 = 100\%$$

nach jedem Intervall verringert sie sich um den jeweiligen Faktor p_x

$$S_x = \prod_{i=1}^x p_i$$

Kaplan-Meier-Verfahren

- Keine fixen Zeitintervalle
- Variable Zeitabschnitte jeweils definiert durch Eintreten eines Zielereignisses
- Detailliertere Information
- Typische Treppenkurven mit unregelmäßigem Bild

Sukzessive Berechnungen für jeden Ereigniszeitpunkt t:

t	Exakter Zeitpunkt des Eintretens eines Zielereignisses
l_t	Lebende Probanden unter Beobachtung unmittelbar vor dem Zeitpunkt t = Überlebende des letzten Ereigniszeitpunktes minus zwischenzeitlich eingetretene Zensierungen
d_t	Zahl der Zielereignisse zum Zeitpunkt t (bei exakter Zeitmessung stets 1, in der Praxis häufig > 1)
q_t	Sterbewahrscheinlichkeit am Zeitpunkt t
p_t	Überlebenswahrscheinlichkeit am Zeitpunkt z
S_t	Kumulative Überlebenswahrscheinlichkeit bis zum Zeitpunkt t

Sterbewahrscheinlichkeit q_x und Überlebenswahrscheinlichkeit p_x im Intervall

$$q_x = \frac{d_t}{l_t} \qquad p_x = 1 - q_x$$

Survival: Kumulative Überlebenswahrscheinlichkeit bis zum Zeitpunkt t

$$S_t = \prod_{i=1}^t p_i$$

Anmerkungen zum Kaplan-Meier-Verfahren:

- Hohe Anforderungen an die Daten: Sterbe und Zensierungsereignisse müssen exakt bekannt sein
- Rechenaufwändiger als das Cutler-Ederer-Verfahren
- Fallstrick bei der Interpretation:
fällt die Überlebenskurve auf Null ab, so bedeutet dies, dass der Patient, der am längsten beobachtet worden ist verstorben ist, das heißt nicht, dass alle Patienten verstorben sind

8 Anhang

8.1 Definitionen

Epidemiologie:

Wissenschaftszweig, der sich mit **der Verteilung von übertragbaren und nicht übertragbaren Erkrankungen und deren Determinanten und Folgen in der Bevölkerung** befasst.

Epidemiology:

„... the study of the distribution and determinants of disease frequency in human populations“
(MacMahon und Pugh 1970)

Inzidenz:

Die Inzidenz zählt **die Neuerkrankungen innerhalb eines definierten Zeitraumes**, und setzt sie in den Bezug zu den am Anfang dieser Beobachtungsperiode nicht kranken Personen

Prävalenz:

Die **Prävalenz** gibt den Anteil kranker Personen zu einem bestimmten Zeitraum (Punktprävalenz) oder den Anteil erkrankter Personen innerhalb eines gewissen Zeitraumes in der Bezugsbevölkerung (Periodenprävalenz) an

Confounding: Klassische Definition:

Ein Confounder ist eine dritte Variable, die mit der interessierenden Exposition assoziiert auftritt und einen unabhängigen Einfluss auf den Effekt ausübt, jedoch keine intermediäre Variable ist.

Confounding: Change-in-estimate Definition (Robins):

Ein Confounder ist eine Variable, deren Berücksichtigung (Kontrolle) zu einem anderen Effektschätzer führt.

Effektmodifikation:

Existieren **in verschiedenen Strata** (Schichten) einer Variablen unterschiedliche Effektschätzer, so spricht man von Effektmodifikation bzw. Interaktion. Die Schichtvariable wird als **Effektmodifikator** (effect **modifier**) bezeichnet. Reine Effektmodifikation führt nicht zu einer Verzerrung des Effektmaßes und gehört damit nicht zu den Fehlern in epidemiologischen Studien.

Direkte Standardisierung:

Die *alters und geschlechts-spezifischen Raten* aus verschiedenen Populationen werden auf eine **Standardpopulation** mit definiertem *Altersaufbau und Geschlechtsverhältnis* angewendet.

Indirekte Standardisierung:

Auf eine Studienpopulation mit *spezifischem Altersaufbau und Geschlechtsverhältnis* werden die *alters- und geschlechtsspezifischen Raten* einer **Vergleichsbevölkerung** angewendet.

Matching:

Beim Matching werden *zueinander passende Personen* unter Erkrankten und Nicht-Erkrankten (in Fall-Kontroll-Studien) bzw. Exponierten und Nicht-Exponierten (in Kohortenstudien) für die Studienpopulation ausgewählt.

Frequency Matching: lediglich gleiche Verteilung der Matchingvariable in zu vergleichenden Gruppen (z.B. Altersgruppen)

Pair Matching: Zuordnung auf individueller Ebene, Proband \Rightarrow 1 - 4 Matching-Partner

Definition Konfidenzintervall:

Das 95%-Konfidenzintervall bedeckt den Parameterwert mit einer Wahrscheinlichkeit von 0.95 (analog 99%-KI-Intervall etc.)

Definition des p-Wertes:

Ein p-Wert ist die Wahrscheinlichkeit, das (vorliegende) Studienergebnis oder ein noch extremeres Ergebnis zu erhalten, wenn die Hypothese H_0 wahr ist.

Bradford-Hill-Kriterien

- | | |
|---|--|
| 1. Stärke der Beziehung | geringe Ausprägung der Assoziation: nur Zufall? |
| 2. Konsistenz der Beziehung | in verschiedenen Studien/Untersuchungen gezeigt? |
| 3. Spezifität des Effektes | klar abgrenzbarer Effekt? |
| 4. Dosis-Wirkungs-Beziehung | höhere Dosis \rightarrow höhere Wirkung? |
| 5. zeitliche Sequenz | Exposition \rightarrow Erkrankung |
| 6. Biologische Plausibilität | Ist die Wirkung plausibel? |
| 7. Kohärenz | Widerspruch zum aktuellen Wissenstand? |
| 8. Experimentelle Evidenz | Überprüfbarkeit der Hypothese |
| 9. Analogie mit ähnlichen Kausalzusammenhängen | gibt es Ähnlichkeiten mit anderen Kausalitäten |

Monitoring:

Kontinuierliche Beobachtung und Erfassung bestimmter Merkmale einer bestimmten Bevölkerungsgruppe

Surveillance:

Systematische fortwährende Sammlung und Analyse von Daten und (kurzfristiges) Feedback an diejenigen, die daraus Rückschlüsse ziehen und darauf reagieren können (sollten): „Daten für Taten“

Ökologische Studie:

Untersuchen den Zusammenhang zwischen Exposition und Erkrankung auf der Ebene aggregierter Populationen (statt Individuen) und Maße (z.B. per capita Rotweinkonsum – KHK-Mortalität auf Länderebene)

8.2 Formeln

Kumulative Inzidenz / Inzidenzdichte:

$$CI = \frac{\text{Zahl Neuerkrankungen}}{\text{Bezugsbevölkerung}} \quad [0-100\% \text{ oder } 0-1]$$

$$ID = \frac{\text{Zahl Neuerkrankungen}}{\text{Personenzeit}} \quad [1/10^5 \text{ Personenjahre}]$$

Abschätzen der ID:

$$ID = \frac{a}{(N - a/2 - D/2 - W/2)(\Delta t)}$$

N = Anzahl der Probanden

a = Fälle

D = konkurrierende Ereignisse (z.B. Tod aus anderem Grund)

W = losses to Follow-Up

Prävalenz:

$$\text{Prävalenz } P = \frac{\text{Zahl bestehender Fälle}}{\text{Gesamtpopulation}}$$

Zusammenhang Prävalenz / Inzidenz:

$$P \approx ID \cdot D \quad (\text{für niedrige Prävalenzen})$$

Zusammenhang kumulative Inzidenz und Inzidenzdichte:

$$CI = 1 - e^{-\sum ID_i \cdot \Delta t}$$

Ist die ID über die Zeit konstant, dann lässt sich die vereinfachen zu

$$CI = 1 - e^{-ID \cdot \Delta t}$$

Für kleine Werte von $CI < 0.1$ gilt

$$CI \approx ID \cdot \Delta t$$

Δt = Beobachtungsdauer

Relatives Risiko / Relative Rate:

$$RR = \frac{CI_1}{CI_0} \quad \text{bzw.} \quad \frac{ID_1}{ID_0}$$

Odds Ratio:

$$OR = \frac{\text{Odds}_1}{\text{Odds}_2} \quad (\approx RR)$$

Risikodifferenz / Attributables Risiko:

$$RD = CI_1 - CI_0 = AR$$

Ratendifferenz / Attributable Rate:

$$RD = ID_1 - ID_0 = AR$$

Number-needed-to-treat:

$$NNT = 1 / RD$$

Attributable Risiko Proportion:

$$ARP = \frac{I_1 - I_0}{I_1} = 1 - \frac{I_0}{I_1}$$
$$ARP = 1 - \frac{1}{RR} = \frac{RR - 1}{RR}$$

Attributable Risiko Proportion in der Gesamtpopulation:

$$PARP = p_e \cdot ARP = p_e \cdot (RR - 1) / RR \approx p_e \cdot (OR - 1) / OR$$

Präventive Fraktion:

$$PF = 1 - RR = 1 - OR$$

Standardisierte Mortalitätsrate:

$$\text{Standardisierte MR} = \frac{\sum N_i \cdot p_i}{\sum N_i} \cdot 10^5$$

N_i = Anzahl der Individuen in der i ten Altersgruppe der Standardpopulation

p_i = Altersspezifische Sterberate der Indexpopulation

Standardisierte Mortalitätsratio:

$$\text{SMR} = \frac{\text{Zahl der beobachteten Fälle in der Indexpopulation}}{\text{Zahl der erwarteten Fälle in der Indexpopulation}}$$

$$\text{SMR} = \frac{\sum r_i}{\sum n_i \cdot P_i}$$

Standardisierte Inzidenz Ratio:

$$\text{SIR} = \frac{\sum r_i}{\sum n_i \cdot P_i}$$

r_i = Anzahl der Sterbefälle in der Indexpopulation

n_i = Anzahl der Individuen in der i ten Altersgruppe der Indexpopulation

P_i = Altersspezif. Sterberate bzw. Inzidenzrate der Standardpopulation

Comparative Mortality Figure

$$\text{CMF} = \frac{\text{MR}_{a_{st}}}{\text{MR}_{sp}} \quad \text{bzw.} \quad \text{CMF} = \frac{\text{MR}_{a_{st}}}{\text{MR}_{b_{st}}}$$

$\text{MR}_{a_{st}}$: standardisierte MR der Indexpopulation a

MR_{sp} : MR der Standardpopulation

$\text{MR}_{b_{st}}$: standardisierte MR der Indexpopulation b

Formel Capture-Recapture-Methode:

$$N = \frac{A \cdot B}{(A \cap B)}$$