

Querschnittsfach

**Epidemiologie,
Medizinische Biometrie und
Medizinische Informatik**

Teil Epidemiologie

(7. Fachsemester)

H.-W. Hense, J. Wellmann, K. Berger

Institut für Epidemiologie und Sozialmedizin

Westfälische Wilhelms-Universität Münster

2012

Dieses Handout erklärt einige wichtige Begriffe und Methoden der Epidemiologie und Biometrie. Es ist weder umfassend noch vollständig und kann ein Lehrbuchstudium nicht ersetzen. Zweck dieses Handouts ist es, kurz gefasste Begriffsbestimmungen bereit zu stellen, beispielsweise um den Aufwand zum Mitschreiben in den Vorlesungen zu reduzieren, die Vorbereitung auf die Übungen zu erleichtern und ferner ein Glossar zum Nachschlagen zur Verfügung zu stellen.

Hinweis für die Medizinische Biometrie

Das internetbasierte Lehr- und Lernsystem JUMBO (Java-unterstützte Münsteraner Biometrie-Oberfläche) bildet die Grundlage für die Übungen in Medizinischer Biometrie. Diese interaktive Oberfläche ist unter dem Link

<http://jumbo.uni-muenster.de/>

erreichbar. Ein komplettes druckbares JUMBO-Skript ist auf der Institutshomepage unter Lehre als Download verfügbar:

www.imib.uni-muenster.de .

Ein wesentlicher Teil der Übungsbeispiele beruht auf den Daten einer klinischen Studie zur Therapie der akuten, myeloischen Leukämie bei Erwachsenen. Eine genaue Beschreibung und Ergebnisse der Studie können in der Publikation (*T. Büchner, W. Hiddemann, B. Wörmann, et al.: Age-Related Risk Profile and Chemotherapy Dose Response in Acute Myeloid Leukemia: A Study by the German Acute Myeloid Leukemia Cooperative Group. Journal of Clinical Oncology 2009 Jan 1;27(1):61-69.*) nachgelesen werden.

Prüfungsrelevantes Basiswissen ist der Stoff der Vorlesungen und der Übungen. Dieser entspricht den folgenden Kapiteln des JUMBO-Systems:

- | | |
|-----------|---|
| Kapitel 1 | Deskriptive Statistik I <ul style="list-style-type: none">• Grundlegende Begriffe• Univariate Verfahren |
| Kapitel 2 | Deskriptive Statistik II <ul style="list-style-type: none">• Schätzung der Überlebensraten nach Kaplan-Meier• Kontingenztafel• Regression und Korrelation |
| Kapitel 3 | Wahrscheinlichkeit, Zufallsvariable, Verteilungen <ul style="list-style-type: none">• Elementare Wahrscheinlichkeitsrechnung• Zufallsvariable• Normalverteilung |
| Kapitel 4 | Statistische Tests, Versuchsplanung <ul style="list-style-type: none">• Grundlagen des statistischen Tests• Spezielle Testverfahren• Versuchsplanung |

Inhaltsverzeichnis

Häufigkeitsmaße in der Epidemiologie	4
Beispielberechnungen	6
Überblick über epidemiologische Studientypen	7
Querschnittsstudie	8
Fall-Kontroll-Studie	9
Die Odds Ratio als Risikomaß in Fall-Kontroll-Studien	10
Kohortenstudie	11
Risikomaße in Kohortenstudien	12
Randomisierte Kontrollierte Studie	14
Diagnostische Tests	15
Bias	17
Konfidenzintervalle und Statistische Tests	20
Altersstandardisierung	23
Kausalität	25
Glossar wichtiger epidemiologischer Begriffe	26

Häufigkeitsmaße in der Epidemiologie

Die Epidemiologie ist eine Wissenschaft, die durch Untersuchungen von Bevölkerungen oder deren Untergruppen versucht, die Häufigkeit, Verteilung und Determinanten von Gesundheit und Krankheit zu ermitteln. Sie beschäftigt sich vor allem mit der Untersuchung eines möglichen Zusammenhanges zwischen Risikofaktoren und Belastungen oder Lebensumständen und dem Auftreten einer oder mehrerer Erkrankungen.

Zur Bestimmung der Häufigkeit einer Erkrankung oder eines Merkmales in einer Bevölkerung werden vor allem zwei Begriffe verwendet:

Unter **Inzidenzrate** (Neuerkrankungsziffer, auch als kumulative Inzidenz oder Absolutes Risiko bezeichnet) versteht man die Zahl neu auftretender Fälle in einer definierten Bevölkerung pro Zeiteinheit (meist pro Jahr). Sie wird aus Vergleichbarkeitsgründen auf eine bestimmte Personenzahl der definierten Bezugsbevölkerung (meist pro 1.000 oder pro 100.000) bezogen. (Siehe genaueres auch unter ‚Risikomaße in Kohortenstudien‘).

Unter **Prävalenz** versteht man den Bestand aller Fälle einer bestimmten Krankheit, eines Merkmales oder eines Risikofaktors zu einem bestimmten Zeitpunkt. Der Bestand wird auf die Gesamtstichprobe oder eine Bevölkerung bezogen. Dieses Häufigkeitsmaß wird als *Punktprävalenz* bezeichnet. Bezieht sich die Erfassung des Bestandes an Krankheitsfällen auf ein Zeitintervall (z.B. bei rekurreierenden Erkrankungen wie Asthma, Kopfschmerzen etc.), spricht man von *Periodenprävalenz*.

Inzidenz und Prävalenz stehen zueinander in Beziehung. Das Verhältnis zwischen Inzidenz und Prävalenz geht aus Abbildung 1 hervor. Die Neuerkrankungsrate (Inzidenz) füllt den „Pool“ der vorhandenen Krankheitsfälle (Prävalenz). Der Aufenthalt im Pool wird durch die Krankheitsdauer bestimmt, die wiederum durch die Heilungsrate bzw. die Sterberate determiniert ist. Das Geschehen lässt sich sehr vereinfacht in die Näherungsformel fassen:

$$\text{Prävalenz} \approx \text{Inzidenz} \text{ mal Krankheitsdauer}$$

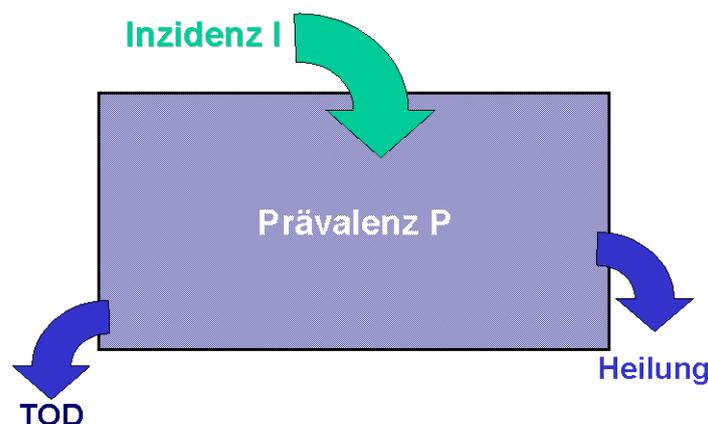
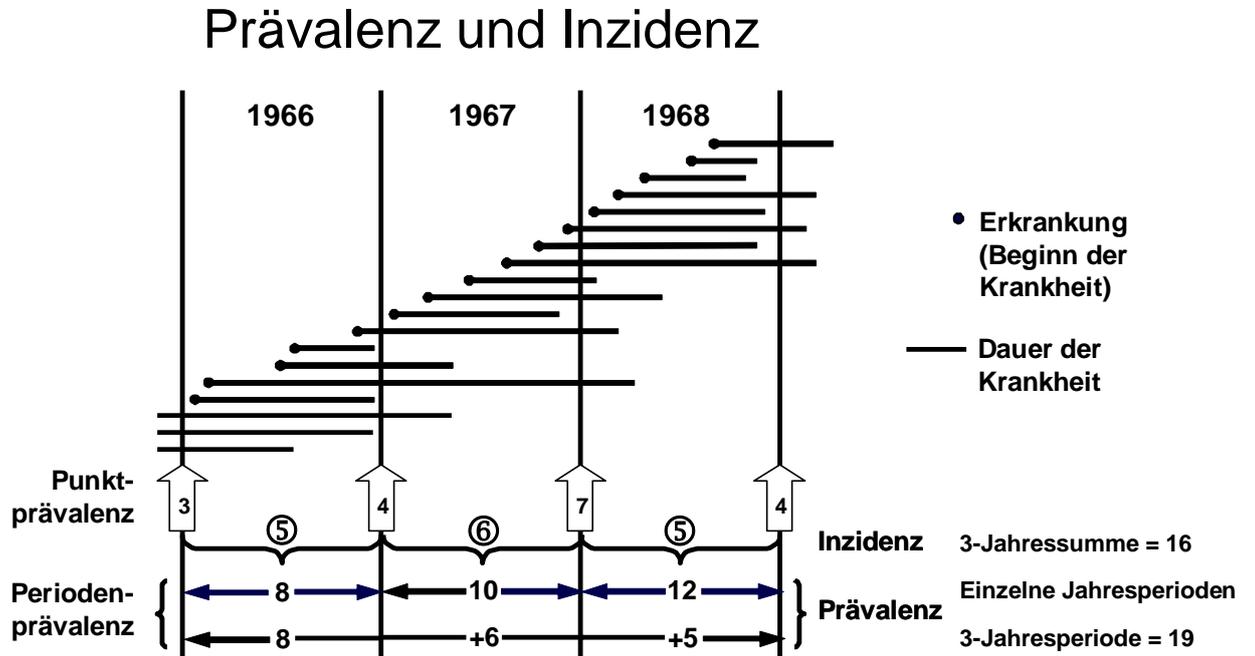


Abbildung 1: Zusammenhang von Inzidenz und Prävalenz

Abbildung 2 gibt Beispiele für die Berechnung von 1–3 Jahres-Inzidenzen und für die Berechnung von Punkt- und Periodenprävalenzen. Jeder Querbalken symbolisiert die Dauer eines Krankheitsfalles.



Grafische Erläuterungen der Begriffe Inzidenz und Prävalenz.
 Jeder Querbalken symbolisiert die Dauer eines Krankheitsfalls.

Aus WHO Cronicle 21, 163 (1976); modifiziert nach Pflanz

Beispielberechnungen

Bei der Berechnung von Raten muss der richtige Zähler (Fälle) auf den richtigen Nenner (Bevölkerung im Risiko) bezogen werden. An zwei Beispielen soll dies verdeutlicht werden:

▪ Beispiel 1:

Eine Stadt hat 100.000 Einwohner (45.000 ♂; 55.000 ♀)

Im Jahre 2000 verstarben 1.000 Personen (600 ♂; 400 ♀)

Im Jahr 2000 erkrankten 50 Personen neu an Lungenkrebs (40 ♂; 10 ♀)

45 an Lungenkrebs Erkrankte verstarben.

Berechne:

- a) rohe Mortalitätsrate = $1.000 / 100.000 = 0,01$; entspricht 10 pro 1.000.
- b) geschlechts-spezifische Mortalität
 - für Männer = $600 / 45.000 = 0,0133$; entspricht 13,3 pro 1.000 Männer.
 - für Frauen = $400 / 55.000 = 0,00727$; entspricht 7,3 pro 1.000 Frauen.
- c) Ursachen-spezifische Mortalität für Lungenkrebs = $45 / 100.000 = 0,00045$; entspricht 0,45 pro 1.000.

▪ Beispiel 2:

Eine Stadt hat 100.000 Einwohner

Im Jahr 1995 verstarben 1.000 Personen (alle Todesursachen)

Die Zahl der im Jahre 1995 an Tuberkulose Erkrankten belief sich auf 300 (200 ♂; 100 ♀)

1995 verstarben außerdem 60 Personen an Tuberkulose, davon 50 Männer.

Berechne:

- a) rohe Mortalitätsrate = $1.000 / 100.000 = 0,01$; entspricht 10 pro 1.000.
- b) die ursachen-spezifische Mortalität für Tuberkulose = $60 / 100.000 = 0,0006$; entspricht 6 pro 10.000 oder 0,6 pro 1.000.
- c) die geschlechts-spezifische Mortalität für Tuberkulose bei Männern = $50 / ?$ bzw. bei Frauen = $10 / ?$ kann mit den vorliegenden Daten nicht berechnet werden.

- Auch das Beispiel zur Altersstandardisierung am Ende dieses Skripts kann zur Ratenberechnung herangezogen werden.

Überblick über epidemiologische Studientypen

▪ **Deskriptive Studien**

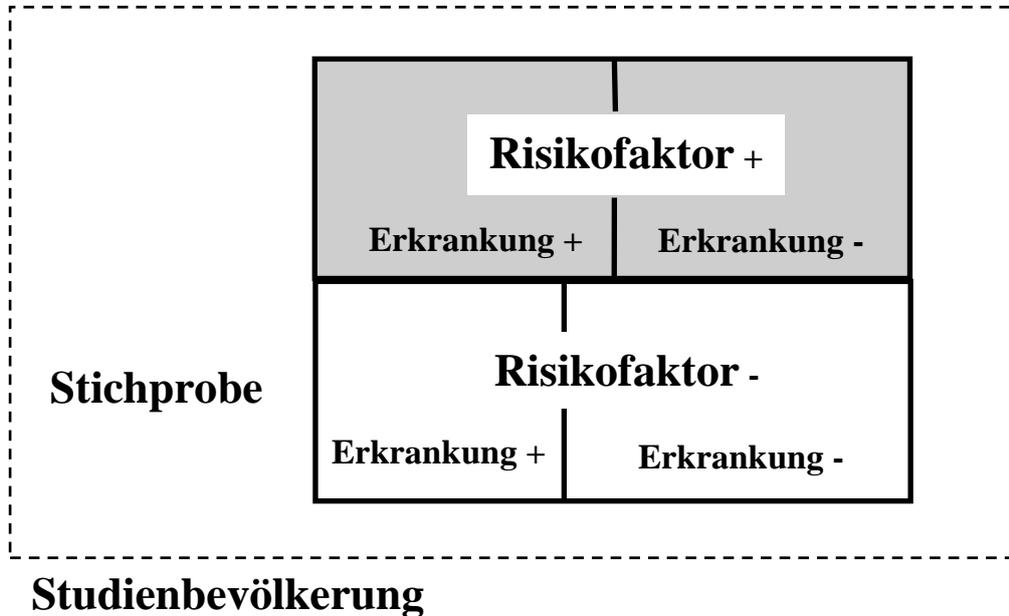
- Studien auf Aggregatdatenebene
Meist ‚ökologisch‘, gelegentlich auch
Korrelationsstudien genannt.
- Studien an Individuen
 - Fall-Berichte
 - Fall-Serien
 - Querschnittsstudien

▪ **Analytische Studien**

- Beobachtungsstudien
 - Fall-Kontroll-Studien
 - Kohortenstudien (auch prospektive oder Längsschnittstudie genannt)
- Interventionsstudien (gelegentlich auch ‚quasi-experimentell‘ genannt)
- Randomisierte, kontrollierte Studien (= RCT, randomised controlled trial)
- Kontrollierte Studien ohne Randomisierung
- Intervention ohne Randomisierung und Kontrollen

Querschnittsstudie

Gegenwart



Vorgehen:

1. Aus einer Bevölkerung (z.B. Wohnbevölkerung, Betriebsangehörige, stationär behandelte Patienten) wird mit geeigneten Verfahren eine Stichprobe gezogen.
2. In dieser Stichprobe werden potentielle Risikofaktoren (z.B. Messungen, Befragungen) und das Vorliegen von Erkrankungen gleichzeitig erhoben.

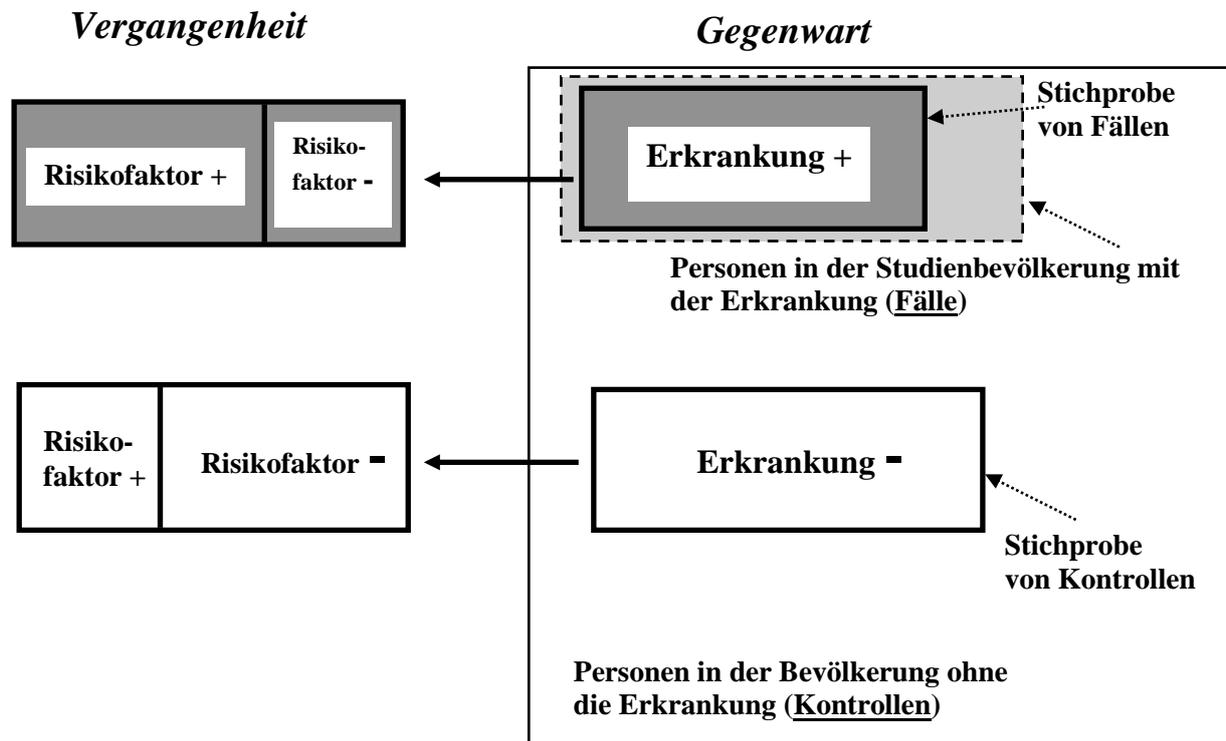
Vorteile:

- Relativ einfach und schnell durchführbar
- Mehrere Erkrankungen (Zielereignisse) können gleichzeitig erfasst werden
- Kontrolle über Stichprobenzusammensetzung und Untersuchungsinstrumente ist gegeben
- Prävalenzen für Risikofaktoren und Krankheiten können berechnet werden

Nachteile:

- Inzidenzen und Relative Risiken können nicht berechnet werden
- Die zeitliche Reihenfolge des Auftretens von Risikofaktor und Erkrankung kann wegen der gleichzeitigen Erfassung nicht bestimmt werden
- Nicht effektiv für selten auftretende Risikofaktoren und Erkrankungen

Fall-Kontroll-Studie



Vorgehen:

1. Aus einer definierten Studienbevölkerung werden (alle oder eine geeignete Stichprobe von) Personen mit der interessierenden Erkrankung (Fälle) identifiziert.
2. Aus der gleichen Studienbevölkerung wird eine geeignete Stichprobe von Personen ohne diese Erkrankung (Kontrollen) gezogen (z.B. Bevölkerungs- oder Krankenhaus-Kontrollen).
3. Die Exposition in der Vergangenheit gegenüber potenziellen Risikofaktoren wird ermittelt.
4. Die Odds (siehe unten) der Risikofaktoren unter den so definierten Fällen und Kontrollen wird dann mit geeigneten Methoden miteinander verglichen.

Vorteile:

- Ermöglicht die Untersuchung selten auftretender Erkrankungen als Fälle.
- Relativ schnell und kostengünstig durchführbar.
- Relative Risiken für einen Risikofaktor können über die Odds Ratio geschätzt werden

Nachteile:

- Nur eine Erkrankung (bei Fällen) kann untersucht werden.
- Weder Prävalenz noch Inzidenz können berechnet werden, da keine Erfassung der Bezugsbevölkerung.
- Zahlreiche Verzerrungsmöglichkeiten müssen beachtet werden (z.B. Verzerrung durch Erinnerung der Exposition, sog. Recall Bias).
- Die korrekte Auswahl der Kontrollen ist schwierig, gelegentlich sehr aufwändig und erfordert viel Erfahrung (wird oft nicht bedacht!).

Die Odds Ratio als Risikomaß in Fall-Kontroll-Studien

Der Begriff stammt aus dem Englischen. Dort spricht man bei einer Erfolgchance von einem Treffer bei zehn Versuchen von einer Trefferwahrscheinlichkeit nicht von 10%, sondern man sagt, die "Odds" für einen Treffer liegen bei 1 zu 9 oder 1/9.

Die Odds Ratio wird meistens im Rahmen von Fall-Kontroll-Studien berechnet und ist dort ein Maß für die Assoziation zwischen Exposition und Erkrankung. Sie entspricht dem Relativen Risiko in Kohortenstudien (s.u.). Es werden die "Odds" für eine Exposition zwischen erkrankten Fällen (Zahl der Exponierten A geteilt durch Zahl der Nicht-Exponierten B) und den entsprechenden Odds unter den nicht-erkrankten Kontrollen (Zahl der Exponierten C geteilt durch Zahl der Nicht-Exponierten D) verglichen. Aus deren Quotient (Ratio) errechnet sich die Odds Ratio wie folgt:

Odds für Exposition unter Fällen (A / B),
 Odds für Exposition unter Kontrollen (C / D)

$$\text{Odds Ratio} = \frac{(A / B)}{(C / D)} = \frac{(A \cdot D)}{(B \cdot C)}$$

Vierfeldertafel zur Berechnung der Odds Ratio in einer Fall-Kontroll-Studie

	Exposition		
	ja	nein	
Fälle	A	B	OR = $\frac{A \cdot D}{B \cdot C}$
Kontrollen	C	D	

Beispiel:

Im Rahmen einer Fall-Kontroll-Studie wurden bei 88 Kindern mit Asthma (Fälle) und 411 Kindern ohne Asthma (Kontrollen) die Rauchgewohnheiten der Mutter (= Exposition) untersucht. Es ergab sich folgende Tabelle:

Mutter raucht +	Mutter raucht -	←	Kind Asthma +	←	Fälle
13	75		88		
Mutter raucht +	Mutter raucht -	←	Kind Asthma -	←	Kontrollen
29	382		411		

Die Odds Ratio errechnet sich als Quotient der "Odds" für eine Exposition bei erkrankten Fällen (13 / 75) und nichterkrankten Kontrollen (29 / 382).

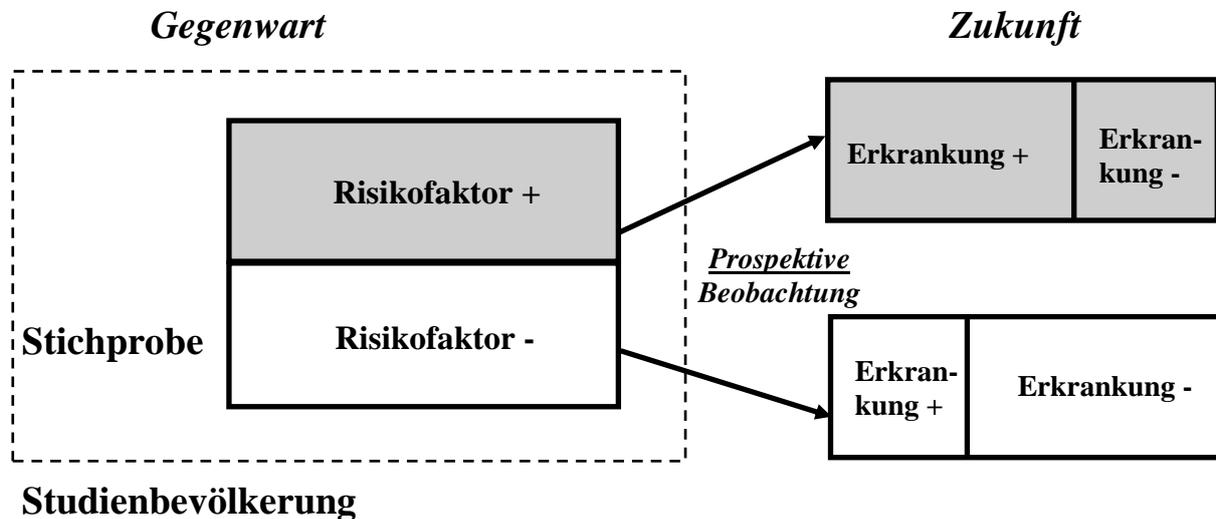
$$\text{Odds Ratio} = \frac{(13/75)}{(29/382)} = \frac{(13 \times 382)}{(29 \times 75)} = 2.28$$

In dieser Studie war somit das Rauchen der Mütter mit einer 2.28-fach erhöhten Odds (Wahrscheinlichkeit) für das Auftreten von Asthma bei ihren Kindern assoziiert.

Man kann auch knapp formulieren:

„Rauchen der Mütter ist mit einem Relativen Risiko von 2.28 assoziiert“.

Kohortenstudie



Vorgehen:

1. Aus einer definierten Studienbevölkerung wird mit geeigneten Mitteln eine Stichprobe gezogen.
2. In dieser Stichprobe werden Risikofaktoren erhoben und Erkrankungen erfasst.
3. Die Stichprobe wird über die Zeit (prospektiv) beobachtet.
4. Am Ende des Beobachtungszeitraumes werden neu aufgetretene Erkrankungen erfasst und zu der Ausgangspräsenz von Risikofaktoren in Beziehung gesetzt.

Vorteile:

- Mehrere Erkrankungen (Zielereignisse) können gleichzeitig erfasst werden
- Krankheits-Inzidenzen und Relative Risiken können berechnet werden
- Die zeitliche Reihenfolge des Auftretens von Risikofaktor und Erkrankung wird berücksichtigt
- Die Zahl der Erkrankungen wächst mit längerer Beobachtung
- Verzerrungsmöglichkeiten (Bias) werden reduziert

Nachteile:

- Kostspielig und langdauernd
- Benötigt oft große Stichproben
- Nicht effektiv für selten auftretende Erkrankungen

Risikomaße in Kohortenstudien

Inzidenzrate:

Anzahl der Neu-Erkrankungen, die während der Beobachtungszeit unter den Studienteilnehmern auftreten. Die Inzidenz ist ein Maß für die Häufigkeit einer Erkrankung.

Man unterscheidet spezifischer die beiden Begriffe kumulative Inzidenz und Inzidenzdichte

Kumulative Inzidenz (auch Absolutes Risiko): Anzahl der Neuerkrankungen bezogen auf eine bestimmte Bevölkerung innerhalb eines definierten Zeitabschnittes. Die Bezugsbevölkerung wird zu Beginn der Studie festgelegt.

Inzidenzdichte: Anzahl der Neuerkrankungen pro beobachteter Personenzzeit. Die Inzidenzdichte bezieht sich nicht mehr auf eine bestimmte Anfangsbevölkerung sondern auf die tatsächlich innerhalb der Studienbevölkerung beobachtete gesamte Personenzzeit, die sich durch die über alle Studienteilnehmer summierten individuellen Zeiten ‚unter Risiko‘ ergibt (z.B. als Personenzjahre). Die individuelle Zeit unter Risiko beginnt mit dem Eintritt einer Person in die Studie und endet entweder mit dem Studienende oder dem Auftreten des zu interessierenden Ereignisses (z.B. der Erkrankung), dem Auftreten eines konkurrierenden Ereignisses (z.B. Tod) bzw. dem vorzeitigen Austritt aus der Studie (z.B. Umzug, Verweigerung).

Relatives Risiko:

Nicht ganz korrekter, aber häufig gebrauchter Begriff: Quotient der Inzidenz unter Exponierten und Nicht-Exponierten. Die Inzidenz kann dabei entweder mit der kumulativen Inzidenz berechnet werden (dann genauer: Risk Ratio) oder mit den Inzidenzdichten (dann genauer: Rate Ratio). Allgemeine Abkürzung RR.

Das RR ist ein Maß für die Assoziation zwischen einer Exposition und einer Erkrankung.

Unter der Bedingung, dass die Inzidenz unter den Exponierten im Zähler und die Inzidenz unter den Nicht-Exponierten im Nenner des Quotienten steht, gilt:

RR = 1.0, wenn die Exposition keinen Einfluss auf die Zielerkrankung hat.

RR > 1.0, wenn die Exposition das Risiko für eine Zielerkrankung erhöht (= Risikofaktor)

RR < 1.0, wenn die Exposition das Risiko für eine Zielerkrankung senkt (= Protektivfaktor)

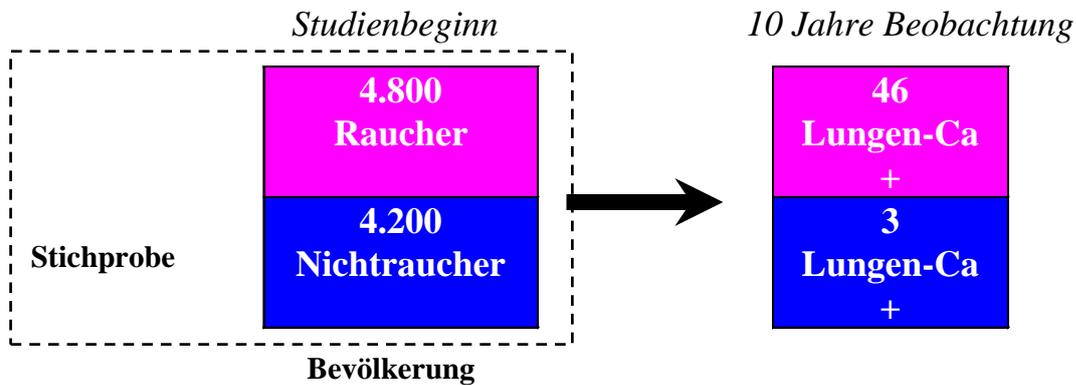
Risikodifferenz (oder Exzess-Risiko, gelegentlich auch Attributables Risiko):

Differenz der Inzidenz unter Exponierten und Nicht-Exponierten.

Maß für die Häufigkeit der Erkrankung unter den Exponierten, die – unter der Voraussetzung, dass keine anderen ursächlichen Faktoren vorliegen – auf die Exposition „zurückzuführen“ ist.

Beispiel:

In einer Kohortenstudie wurden 4800 Raucher und 4200 Nichtraucher über 10 Jahren bezüglich des Auftretens von Lungenkrebs beobachtet.



Inzidenz bei Rauchern (I_R): $I_R = 46 / 4800 = 0.0096$

d.h., 96 neue Fälle von Lungenkrebs auf 10.000 Raucher in 10 Jahren.

Inzidenz bei Nichtrauchern (I_{NR}): $I_{NR} = 3 / 4200 = 0.0007$

d.h., 7 neue Fälle von Lungenkrebs auf 10.000 Nichtraucher in 10 Jahren.

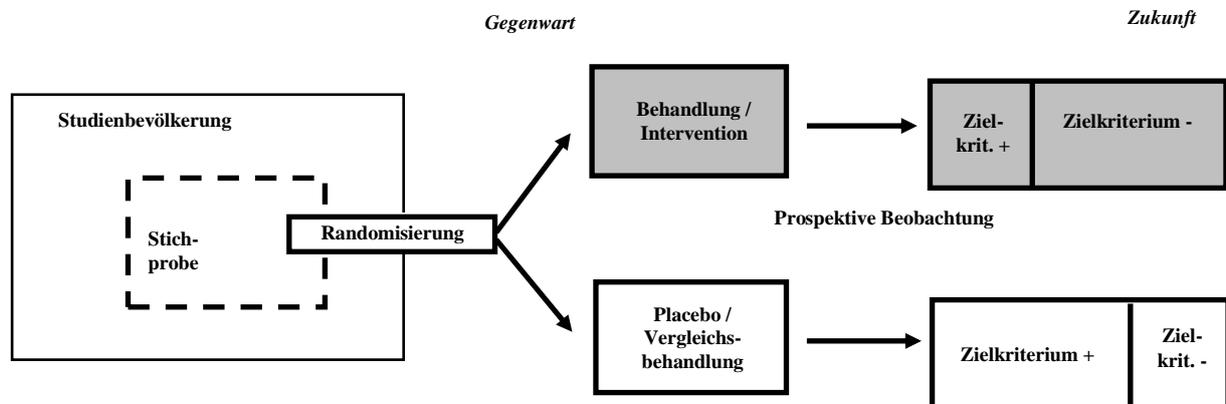
Relatives Risiko: $I_R / I_{NR} = 0.0096 / 0.0007 = 13.7$

d.h. Raucher erkranken über 10 Jahre 13.7-mal häufiger an Lungenkrebs als Nichtraucher.

Risikodifferenz (Attributables Risiko): $I_R - I_{NR} = (0.0096 - 0.0007) = 0.0089$

d.h., 89 der 96 Lungenkrebsfälle unter Rauchern sind auf das Rauchen „zurückführbar“.

Randomisierte Kontrollierte Studie



Vorgehen:

1. Aus einer Studienbevölkerung wird eine definierte Stichprobe gezogen (meistens Patienten mit einer bestimmten Erkrankung oder Risikokonstellation).
2. Die Personen der Stichprobe werden nach einem Zufallsverfahren (Randomisierung) auf eine Interventions- bzw. Placebo/Vergleichsgruppe aufgeteilt.
3. Die Beobachtung über die Zeit (prospektiv) und Vergleich der Häufigkeit des Auftretens zuvor definierter Zielkriterien (z.B. Sterblichkeit, Neuerkrankungen) in den beiden Gruppen mit geeigneten Methoden.

Vorteile:

- Gibt die stärkste epidemiologische Evidenz für die Kausalität einer Beziehung.
- Erlaubt die Schätzung der Wirkung einer Intervention/Behandlung auf das Zielkriterium.
- Randomisierung vermeidet weitgehend Bias.
- Blinding (einfach- oder besser: doppelt-blind) reduziert Einfluss von Behandler bzw. Behandeltem auf die Zielkriterien.

Nachteile:

- Zeit- und kostenaufwendig
- Nicht geeignet bei seltenen Zielkriterien.
- Untersucht nur eine oder wenige Interventionen.
- Externe Validität (Generalisierbarkeit) kann limitiert sein (abhängig von den Einschlusskriterien).
- Aus ethischen Gründen nicht bei allen Fragestellungen durchführbar.

Risikomaße:

wie in Kohortenstudien, also Relatives Risiko (RR) oder Risikodifferenz. Oft auch:

„Absolute Risikoreduktion“ = Risikodifferenz zwischen Interventions- und Vergleichsgruppe

„Relative Risikoreduktion“ = $[1 - RR]$, wird in Prozent ausgedrückt.

(z.B. wenn $RR = 0.67$, dann $[1 - RR] = 0.33$; in Worten: relative Risikoreduktion = 33%).

Diagnostische Tests

Diagnostische Tests zielen darauf ab, möglichst wirkungsvoll **Kranke** von **Gesunden** zu unterscheiden. Hierzu werden **Diagnoseverfahren** oder **Tests** verwendet, deren Eigenschaften und Güte es zu bestimmen gilt oder die bekannt sind. In Abhängigkeit von diesen Eigenschaften kann ein diagnostischer Test dazu beitragen, die nachfolgend erforderlichen klinischen Schritte – weitere Diagnostik, Therapie oder Intervention – in angemessener Weise zu lenken.

Die Prinzipien der diagnostischen Entscheidungsfindung beinhalten eine Abschätzung der Vortest-Wahrscheinlichkeit für die betreffende Erkrankung und die verlässliche Information über die jeweiligen Eigenschaften und die Diskriminationskraft eines bestimmten Tests. Die mathematischen Grundlagen der nachfolgenden Ausführungen wurden bereits frühzeitig in Bayes' Theorem (Thomas Bayes, 1702–1761) formuliert.

Schematisch lässt sich die diagnostische Situation in einer Vier-Felder-Tafel darstellen:

Ergebnisse des diagnostischen Tests	Krankheit vorhanden	Krankheit nicht vorhanden	
	Test positiv	A Richtig Positive (RP)	B Falsch Positive (FP)
Test negativ	C Falsch Negative (FN)	D Richtig Negative (RN)	C + D
	A + C	B + D	A + B + C + D

Die Eigenschaften des Tests lassen sich wie folgt beschreiben:

Sensitivität = $A / (A + C)$: Anteil der RP an allen Erkrankten
Spezifität = $D / (B + D)$: Anteil der RN an allen Nicht-Erkrankten

Sensitivität und Spezifität sind oft weitgehend konstante Eigenschaften eines Tests und unabhängig von der Situation der Anwendung.

Positiver Prädiktivwert = $A / (A + B)$: Anteil der RP an allen Test-Positiven
Negativer Prädiktivwert = $D / (C + D)$: Anteil der RN an allen Test-Negativen

Positiver und Negativer Prädiktivwert sind keine konstanten Charakteristika eines Tests, sondern ändern sich in Abhängigkeit von der Untersuchungssituation. Dabei sind die Prädiktivwerte abhängig von der Häufigkeit der Erkrankten unter allen Untersuchten, das heißt, von $(A + C) / (A + B + C + D)$ (diesen Wert kann man als Prävalenz der Erkrankung unter allen Untersuchten auffassen, er wird auch als die so genannte Vor-Test-Wahrscheinlichkeit für die Erkrankung bezeichnet). Diese komplexen Zusammenhänge werden durch Bayes' Theorem quantitativ beschrieben.

Vereinfacht gilt: Je höher Prävalenz/Vor-Test-Wahrscheinlichkeit einer Krankheit, desto höher ist auch der Positive Prädiktivwert eines Tests mit gegebenen Eigenschaften (also: bei gegebener Sensitivität und Spezifität). Umgekehrt, je geringer die Prävalenz/Vor-Test-Wahrscheinlichkeit einer Krankheit, desto höher ist auch der Negative Prädiktivwert eines Tests mit gegebenen Eigenschaften.

Dies bedeutet für die klinische Entscheidung, dass die Wahrscheinlichkeit für das Vorliegen einer Erkrankung (bzw. ihre Abwesenheit) immer von der Zusammensetzung des Patientengutes abhängt. Der gleiche Test kann in Abhängigkeit davon, ob er beim Screening einer in sehr großen Teilen nicht erkrankten Bevölkerung eingesetzt wird oder aber in einer fachärztlichen Praxis sehr unterschiedliche Prädiktivwerte aufweisen.

Nur unter genauer Kenntnis dieser Zusammenhänge sowie bei Vorliegen von Informationen über die Sensitivität und Spezifität des jeweils eingesetzten Tests kann eine angemessene diagnostische Entscheidung gefällt werden.

Bias

Bias (deutsch: Verzerrung) ist ein Vorgang, der bei einer Studiendurchführung bzw. -auswertung zu einer systematischen Abweichung von den wirklichen Werten führt.

Bei der Beurteilung der Untersuchungsergebnisse von (epidemiologischen) Studien stellt sich immer die Frage, inwieweit die Untersuchung tatsächlich das misst, was sie zu messen beabsichtigt (Validität). Es muss deshalb bei jeder im Rahmen einer Analyse gefundenen statistisch signifikanten Assoziation (gleichgültig, ob diese positiv oder negativ ist) überlegt werden, ob der gefundene Zusammenhang nicht durch Bias bedingt sein kann.

Wichtige Arten von Verzerrung in der Epidemiologie sind:

- **Selektionsbias**
- **Informationsbias**
- **Confounding**

- **Selektionsbias**

Bei der Rekrutierung der Untersuchungsteilnehmer für eine Studie ist es von größter Wichtigkeit, dass bestimmte Personengruppen einer zu untersuchenden Population in der Studie nicht unter- oder überrepräsentiert sind.

Beispiel:

Soll an Schulen die Asthma-Prävalenz bei Jugendlichen bestimmt werden, sind die Schüler/innen, die bei einer ersten Untersuchung/Befragung abwesend sind, unbedingt zu einem späteren Termin zu untersuchen/befragen. Gerade diese Schüler/innen könnten wegen asthmatischer Beschwerden abwesend sein; die tatsächliche Asthma-Prävalenz bei jugendlichen Schülern würde andernfalls unterschätzt.

- **Informationsbias**

Nachdem die Studienteilnehmer ausgewählt sind, müssen von den Probanden die zu vergleichenden Informationen (z.B. zur Exposition oder Erkrankung) gewonnen werden. Durch systematische Unterschiede beim Antwortverhalten der Probanden bzw. bei der Erhebung von Daten und Informationen in den zu vergleichenden Studiengruppen kann ein Informationsbias resultieren. Diese Verzerrung stellt z.B. bei Fall-Kontroll-Studien ein Problem dar.

Beispiel 1:

In einer Fall-Kontroll-Studie wurden die Erkrankten (= Fälle) möglicherweise schon mehrmals vor dem Beginn der Studie zu bestimmten Expositionen befragt (z.B. Asbestexposition bei Lungenkrebspatienten), während die nicht an dieser Krankheit leidenden Kontrollprobanden vielleicht erstmalig im Rahmen der Studie befragt werden. Dies kann dazu führen, dass eine seltene oder lange zurückliegende Exposition von den Kontrollen weniger häufig angegeben bzw. erinnert wird. Aufgrund dieses so genannten „Recall Bias“, der eine spezielle Form des Informationsbias ist, kann sich ein fälschlicherweise erhöhtes Relatives Risiko (→ Odds Ratio) für die Fälle ergeben.

Beispiel 2:

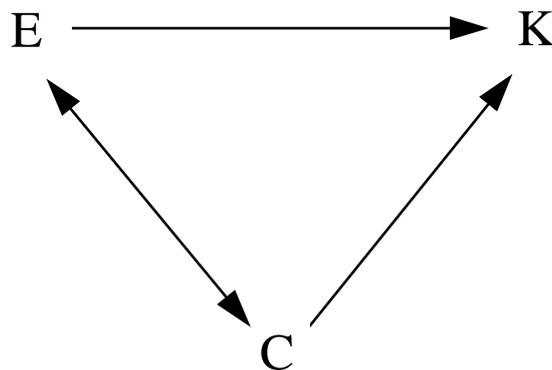
Bestimmte Faktoren sind hinsichtlich ihrer gesundheitlichen Auswirkungen bekannt oder sie sind sozial erwünscht bzw. unerwünscht. Bei Befragungen kann dies zu einem Bias führen, wenn die Befragten ihre wahre Exposition ‚beschönigen‘ wollen. So kann das Rauchen eventuell gänzlich geleugnet oder als zu gering angegeben werden, ähnliches gilt für Angaben zum Alkoholkonsum. Auch Selbstangaben zum Gewicht sind meist gebiast (zu gering...).

- **Confounder** (lat.: confundere = vermischen, vermengen)

Confounding ist die Verzerrung eines Schätzwertes für eine epidemiologische Maßzahl, die aufgrund des Einflusses eines zusätzlichen Faktors („Confounder“) entsteht. In seiner einfachsten Form lässt sich Confounding als eine „Vermischung mehrerer Faktoren, die einen Einfluss auf die Entwicklung einer Krankheit haben (mixing of effects)“, beschreiben.

Ein Confounder muss folgende 3 Bedingungen erfüllen:

- ❶ Der Faktor muss selbst ein Risikofaktor oder ein präventiver Faktor in Bezug auf die untersuchte Krankheit sein
- ❷ Der Faktor muss mit der untersuchten Exposition assoziiert sein.

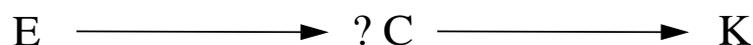


Beziehung zwischen einer Exposition (**E**), einer Krankheit (**K**)
und einem Confounder (**C**)

Beispiel:

Wenn man die mögliche protektive Wirkung von Alkohol (E) auf die Entwicklung der koronaren Herzkrankheit (K) untersucht, muss man den Faktor Rauchen (R) als möglichen Confounder erwägen, da Rauchen ein eigenständiger Risikofaktor der KHK ist und meist auch positiv mit dem Alkoholkonsum assoziiert ist. Wenn man in der Datenanalyse nicht für den Confounder Rauchen „kontrolliert“, könnte sich ein möglicher protektiver Effekt von Alkohol auf die KHK nur vermindert oder gar nicht abbilden.

- ❸ Der Faktor darf nicht lediglich ein Zwischenschritt in einer Kausalkette sein.



Beziehung zwischen einer Exposition (**E**), einer Krankheit (**K**) und einem vermeintlichen Confounder (**? C**), der hier allerdings einen pathophysiologischen Zwischenfaktor in einer kausalen Kette darstellt und damit kein Confounder ist.

Beispiel:

Je mehr krebserzeugende Polyzyklische Aromatische Kohlenwasserstoffe (PAHs) beim Rauchen eingeatmet werden, desto häufiger ist eine Anlagerung von PAH an Kern-DNA im Lungengewebe nachweisbar (sog. DNA-Addukte). Bei der Untersuchung des Zusammenhanges zwischen Rauchen und Lungenkrebs sollte die Menge von PAH-DNA-Addukten nicht als Confounder berücksichtigt werden, da die Addukte höchstwahrscheinlich lediglich ein Zwischenschritt in der kausalen Kette „Rauchen → Lungenkrebs“ darstellen. Eine korrekte Datenanalyse ist deshalb oft schwierig und erfordert ein gutes Verständnis der biomedizinischen Krankheitsgrundlagen!

Konfidenzintervalle und Statistische Tests

Die Epidemiologie ist eine empirische Wissenschaft, die versucht, Gesundheitseffekte in Bevölkerungen zu quantifizieren. Eine epidemiologische Studie wird (im Gegensatz zum häufig replizierbaren Experiment) nur einmal durchgeführt und das in dieser Studie ermittelte Effektmaß ist eine einmalige *Schätzung* des wahren Wertes.

Konfidenzintervall

Das Konfidenzintervall ist ein statistisch bestimmtes Maß für die Präzision, mit der eine Studie bestimmte Effekte (z.B. Risiken, Raten, Prävalenzen) geschätzt hat.

Das Konfidenzintervall besagt theoretisch, dass bei einer 100-mal mit gleicher Methodik durchgeführten Studie, der Schätzwert (für eine Assoziation oder eine Häufigkeit) wahrscheinlich in 95 Studien innerhalb des vom 95%-Konfidenzintervall umschlossenen Wertebereichs liegen würde, in 5 Studien aber würde er außerhalb dieses Bereiches liegen können.

Das Konfidenzintervall umfasst den Bereich möglicher Schätzwerte, der mit den beobachteten Daten *im Sinne bestimmter statistischer Annahmen* verträglich ist. Man führt diese **Intervall-Schätzung** aus, um bei der in der Regel einmaligen Durchführung einer Studie der Zufälligkeit gerecht zu werden, mit welcher der einzelne Schätzwert (**Punkt-Schätzer**) für die epidemiologischen Häufigkeits- oder Assoziationsmaße aus theoretisch unendlich vielen Möglichkeiten hervorgeht.

Die Weite des Konfidenzintervalls hängt ab von der Variabilität der beobachteten Daten:
Vergrößerung der Stichprobe (= Reduktion der zufallsbedingten Varianz),
Verbesserung der Messgenauigkeit (= Reduktion der Messfehlerbedingten Varianz)
verkleinern das Konfidenzintervall.

Unter der unabdingbaren Voraussetzung, dass die Studie ohne systematische Verzerrungen durchgeführt wurde, kann man im Allgemeinen davon ausgehen, dass der in der untersuchten Population geltende „wahre“ Wert mit 95%iger Wahrscheinlichkeit in dem 95%-Konfidenzintervall liegt.

Beispiel:

Das Relative Risiko (RR) für Lungenkrebs bei Passivrauchern im Vergleich zu Nicht-Rauch-exponierten wurde in einer Studie geschätzt als:

RR = 1.35 mit einem 95%-Konfidenzintervall von 1.20 bis 1.53.

Das heißt, in dieser (als valide vorausgesetzten!) Studie erlauben Teilnehmerzahl, Messgenauigkeit und Häufigkeit des Ereignisses Lungenkrebs die Aussage, dass das Relative Risiko für Lungenkrebs bei Passivrauchern mit einer Wahrscheinlichkeit von 95% zwischen den Werten 1.20 und 1.53 liegt.

Statistische Tests

Statistisches Testen konzentriert sich auf eine *Entscheidung* darüber, ob eine bestimmte Studienhypothese mit einer vorgegebenen Irrtumswahrscheinlichkeit abgelehnt werden kann oder nicht. Dabei nimmt die sogenannte Null-Hypothese an, dass zwischen zwei Faktoren kein Zusammenhang besteht.

Der *p-Wert* ist das gebräuchliche statistische Maß bei der Hypothesen-Testung und beschreibt die Wahrscheinlichkeit, mit der eine beobachtete Assoziation zwischen zwei Faktoren in einer bestimmten Stichprobe auftritt, wenn die Null-Hypothese wahr ist. p-Werte werden anhand von statistischen Modellen berechnet, von denen man annimmt, dass sie die Beziehungen zwischen beobachteten Daten adäquat beschreiben. Der p-Wert ist eine kontinuierliche statistische Größe, die zwischen 0 und 1 liegt.

Niedrige p-Werte indizieren also eine geringe Wahrscheinlichkeit dafür, dass die beobachteten Daten mit der Null-Hypothese in Einklang sind. Hohe p-Werte indizieren, dass die beobachtete Assoziation unter der Null-Hypothese häufig beobachtet wird, die Ergebnisse also damit in Einklang sind.

Die Null-Hypothese wird nur dann zugunsten der Alternativ-Hypothese abgelehnt, wenn ein bestimmter, zuvor festgelegter Grenzwert (die Irrtumswahrscheinlichkeit α) unterschritten wird. Dieser beträgt **konventionell 5%, was einem p-Wert von < 0.05 entspricht**.

Ein Studienresultat wird dann als **statistisch signifikant** bezeichnet, wenn der durchgeführte statistische Test einen p-Wert < 0.05 aufweist. Statistisch signifikant heißt in diesem Fall also, dass das Resultat als nicht mehr mit der Null-Hypothese („kein Unterschied“) kompatibel eingestuft und die Alternativ-Hypothese („Unterschied vorhanden“) angenommen wird.

Beispiel:

Das Relative Risiko für Lungenkrebs bei Passivrauchern wurde in einer Studie geschätzt als: $RR = 1.35$ mit einem P-Wert von $p = 0,075$.

Bei der konventionellen Irrtumswahrscheinlichkeit von 5% wird die Null-Hypothese nicht verworfen, da $p = 0.075 > \alpha = 0.05$! Man sagt, das Relative Risiko ist auf dem 5%-Niveau statistisch nicht signifikant.

Im Fall des Testens werden vielfältig informative Studienergebnisse auf eine dichotome Entscheidung für oder gegen eine statistische ‘Signifikanz’ reduziert. Beachte dabei: statistische Signifikanz ist nicht gleich klinischer Relevanz!

Zusammenhang zwischen Schätzen und Testen

Auf Grund der statistischen Herleitung besteht eine direkte Beziehung zwischen dem Konfidenzgrad (definiert als $[1 - \text{Irrtumswahrscheinlichkeit } \alpha]$) bei der Intervall-Schätzung und dem α -Wert beim Hypothesen-Testen.

Ist $\alpha = 0.05$ bzw. die Irrtumswahrscheinlichkeit 5%, so ergibt sich der Vertrauensgrad des Konfidenzintervalls als $(1 - 0.05) = 0.95$ oder 95%. Wenn das 95%-Konfidenzintervall den in der Null-Hypothese angenommenen Wert (z.B. ein Relatives Risiko von 1.0 oder eine Risikodifferenz von 0) *nicht* enthält, dann ist dies gleich bedeutend mit der konventionellen Ablehnung der Null-Hypothese mit einer Irrtumswahrscheinlichkeit von 5%.

Epidemiologische Studien versuchen, die Richtung und Größenordnung gesundheitsrelevanter Zusammenhänge zu schätzen. Dazu gehört im Studienergebnis die Angabe, wie genau die Schätzung erfolgen konnte. Valide durchgeführte Studien mit geringerer Teilnehmerzahl oder der Untersuchung seltener Krankheitsereignisse berichten häufig weite Konfidenzintervalle, die im konventionellen Sinne nicht signifikant sind. Dennoch können dabei einzigartige und wichtige Erkenntnisse gewonnen werden und die Studienresultate bieten gelegentlich die aktuell beste verfügbare Evidenz zu einem bestimmten Thema. Es wäre unsinnig, solche Ergebnisse als ‘nicht signifikant’ zu verwerfen.

Altersstandardisierung

Die Altersstandardisierung ist eine analytische Methode, die einen unverzerrten Vergleich von Populationen mit unterschiedlicher Altersstruktur ermöglicht.

Das Lebensalter ist eine wesentliche Determinante vieler Krankheiten und damit ein potentieller Confounder (s. *BIAS*) bei Gruppen- oder Bevölkerungsvergleichen: ein valider Vergleich von epidemiologischen Häufigkeitsmaßen ist deshalb nur bei gleicher Altersverteilung der Studienpopulationen möglich. Da dies im Allgemeinen aber nicht der Fall ist, werden die Häufigkeitsmaße der zu vergleichenden Populationen auf die Altersverteilung einer externen Referenzpopulation "standardisiert", um altersbedingte Verzerrungen zu eliminieren.

Beispiel:

Anzahl der Einwohner, Todesfälle und Mortalitätsrate (auf 10 000 Einwohner und Jahr), nach Altersgruppen und Bundesstaat der U.S.A., 1974.

Altersgruppe	Florida			Alaska			U.S.A.
	Einwohner	Todesfälle	Mortalitätsrate	Einwohner	Todesfälle	Mortalitätsrate	Einwohner in %
<45	5 000 000	2 000	4	300 000	300	10	69.2
45–65	1 700 000	8 000	47	45 000	300	67	20.4
>65	1 300 000	70 000	538	5 000	400	800	10.4
Gesamt	8 000 000	80 000	100	350 000	1 000	29	100.0

Die 'rohen' (also nicht altersstandardisierten!) Gesamtmortalitätsraten betragen:

Florida: $80\,000 / 8\,000\,000 = 0,01$ / Jahr, oder $100 / 10\,000$ Einwohner und Jahr;

Alaska: $1\,000 / 350\,000 = 0,0029$ / Jahr oder $29 / 10\,000$ Einwohner und Jahr.

Die **rohe Gesamtmortalitätsrate** in Florida war *höher* als in Alaska, obwohl in Florida die Mortalität in jeder der drei Altersgruppen (**altersspezifische Mortalitätsraten**) *niedriger* war als in Alaska.

Eine direkte Altersstandardisierung (hier auf die Altersstruktur der Bevölkerung der U.S.A. als 'Standard') erfolgt, indem man die altersspezifischen Mortalitätsraten von Florida und von Alaska mit dem Prozentanteil an der Standardbevölkerung multipliziert und sodann über alle Altersgruppen summiert. Dies ergibt die **altersstandardisierten Gesamtmortalitätsraten**:

Florida: $0.692 \cdot 4 + 0.204 \cdot 47 + 0.104 \cdot 538 = 69$ (auf 10.000 Einwohner und Jahr)

Alaska: $0.692 \cdot 10 + 0.204 \cdot 67 + 0.104 \cdot 800 = 104$ (auf 10.000 Einwohner und Jahr)

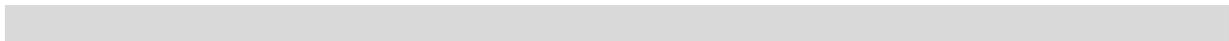
Die altersstandardisierte Gesamtmortalitätsrate, welche die völlig unterschiedliche Altersstruktur der beiden Bundesstaaten eliminiert, ist in Alaska also *höher* als in Florida!!

Beachte:

Rohe Raten sind geeignet, die reale Größe eines Gesundheitsproblems in einer Bevölkerung zu beschreiben. Im obigen Beispiel bedeutet dies, dass in Florida 100 von 10 000 Einwohnern jährlich versterben, während dies in Alaska nur bei 29 Personen der Fall ist. Die Gesamtsterblichkeit über alle Altersgruppen ist in Florida deshalb höher, weil dort vor allem viele Ältere leben. Diese Zahl ist gesundheitsplanerisch bedeutsam.

Die altersspezifischen Raten zeigen die Sterblichkeit in einem engen Alterbereich an. Im Beispiel wird deutlich, dass in jeder der drei Altersgruppen die Sterblichkeit in Florida niedriger liegt als in Alaska. Altersspezifische Raten sind der *informativste* Weg zur Beschreibung der Raten in einer Bevölkerung, da sie „echte“ Bevölkerungsdaten und weitgehend frei von Alterseinflüssen sind. Allerdings ist ihre Darstellung häufig sehr umständlich und unübersichtlich.

Will man in einem Bevölkerungsvergleich den Einfluss verschiedener Altersstrukturen beseitigen und dennoch über alle Altersgruppen eine übersichtliche Gesamtaussage machen, so benutzt man altersstandardisierte Raten. Altersstandardisierte Raten sind „unechte“ Raten, da sie von der Struktur der jeweils gewählten Standardpopulation abhängen! Sie geben in unserem Beispiel Auskunft darüber, wie viele Sterbefälle insgesamt zu erwarten *wären*, wenn die Gesamtbevölkerung der USA sterberaten wie Florida oder wie Alaska aufweisen würde.



Kausalität

Die Beziehung zwischen Ursachen und den Effekten, die sie hervorrufen. Die Epidemiologie beschäftigt sich intensiv mit Kausalität, wobei verschiedene Arten von Ursachen unterschieden werden können (z.B. notwendige und hinreichende). Die epidemiologische Evidenz per se kann Kausalität nicht belegen, aber sie bietet oftmals wichtige Grundlagen.

Wie bereits ausführlich dargelegt beschäftigt sich die Epidemiologie mit der Untersuchung eines möglichen Zusammenhanges zwischen Exposition und Erkrankung. Der Nachweis eines statistischen Zusammenhanges zwischen einer Exposition, z.B. einem Risikofaktor, und einer Erkrankung bedeutet jedoch noch nicht, dass dieser Faktor für die Entstehung der Erkrankung ursächlich ist (**kausaler Zusammenhang**). Um Entscheidungshilfen hinsichtlich der Beurteilung der möglichen Kausalität einer Exposition für eine Erkrankung bereitzustellen, wurden verschiedene **Kausalitätskriterien** definiert:

- *Experimentelle Evidenz*: Ist der Zusammenhang in Studien der ‚experimentellen Epidemiologie‘ (Interventionsstudien) bestätigt worden?
- *Stärke der Beziehung*: Besteht ein großer Zusammenhang zwischen Exposition und Erkrankung? Stärkere Beziehungen sind weniger wahrscheinlich nur durch Zufall erklärbar.
- *Konsistenz der Beziehung*: Wurde der gefundene Zusammenhang bereits in unterschiedlichen Studien, in verschiedenen Studienbevölkerungen und unter unterschiedlichen Bedingungen nachgewiesen?
- *Spezifität des Effektes*: Kann der gefundene Zusammenhang klar abgegrenzt werden?
- *Zeitliche Sequenz*: Liegt die Exposition nachweislich bereits vor der Erkrankung vor?
- *Dosis-Wirkungs-Beziehung*: Findet sich bei einer höheren Dosis der Exposition auch eine höhere Wirkung auf die Erkrankungshäufigkeit?
- *Biologische Plausibilität*: Ist der gefundene Zusammenhang biologisch plausibel?
- *Kohärenz*: Steht der gefundene Zusammenhang im Einklang mit dem aktuellen Wissenstand, z.B. zur Pathogenese der Erkrankung?
- *Analogie mit ähnlichen Kausalzusammenhängen*: Ist der gefundene Zusammenhang ähnlich zu anderen vergleichbaren Kausalzusammenhängen?

Es ist darauf hinzuweisen, dass keines der vorgestellten Kausalitätskriterien allein hinreichend ist und die Kriterien nicht unkritisch angewandt werden sollten. Eine Bewertung der Kausalität sollte im Zusammenhang aller zur Verfügung stehenden wissenschaftlichen Erkenntnisse erfolgen, da immer wieder Ausnahmen zu den genannten Kriterien auftreten können (mit Ausnahme der zeitlichen Sequenz...).

Glossar wichtiger epidemiologischer Begriffe

Absolutes Risiko:

Die Wahrscheinlichkeit, mit der ein Ereignis, eine Erkrankung oder der Tod während eines bestimmten Zeitraumes in einer Studienpopulation eintritt.

Aggregatdaten:

Aggregatdaten sind Daten, die durch Zusammenfassung (Aggregation) von Individualdaten entstehen, z.B. durch Summierung oder Bildung von Durchschnittswerten. Beispiel: Durchschnittlicher Alkoholkonsum pro Kopf und Jahr in Deutschland.

Analytische Epidemiologie.

Teil der Epidemiologie, welcher versucht, durch Überprüfung von Studienhypothesen neue Hinweise auf Krankheitsursachen zu erhalten.

Attributables Risiko:

Der Anteil einer Erkrankung unter den Exponierten, welcher der Exposition (kausal) zugeordnet werden kann. Entspricht der Differenz der absoluten Risiken (z.B. kumulative Inzidenz oder Mortalität) unter Exponierten und Nicht-Exponierten. Dabei wird angenommen, dass andere Ursachen unter Exponierten und Nicht-Exponierten die gleiche Wirkung haben (Fehlen von Interaktionen).

Bias:

Systematische, unbewusste Verzerrung. Das englische Wort Bias (sprich: beies) hat sich im Deutschen eingebürgert. Nach Ursprung der Verzerrung werden verschiedene Arten unterschieden, die wichtigsten sind:

- Selektionsbias (z.B. Responder Bias, also selektive Teilnahme etc.)
- Informationsbias (z.B. Recall Bias, also selektive Erinnerung; Messfehler etc.)
- Confounding (s. unten)

Confounding:

Der Begriff kommt vom lateinischen Verb confundere = vermischen. Es handelt sich um eine Verzerrung, die dadurch hervorgerufen wird, dass das Maß (z.B. Relatives Risiko) des Effektes einer Exposition auf ein Krankheitsrisiko verfälscht wird, weil ein Zusammenhang zwischen der Exposition und einem (oder mehreren) anderen Faktor(en) vorliegt, der (die) selbst das Krankheitsrisiko beeinflusst. Ein Confounder wird im Deutschen gelegentlich auch als Störvariable bezeichnet.

Deskriptive Epidemiologie:

Befasst sich mit der Beschreibung der Häufigkeit von Erkrankungen oder Gesundheitsstörungen und deren Verteilung in der Bevölkerung.

Experimentelle Epidemiologie:

Teil der Epidemiologie, der sich mit der Testung eines experimentellen, kontrollierbaren Faktors befasst: Intervention im Rahmen kontrollierter Studien, randomisierte klinische Studien, bevölkerungsbezogene Interventionsstudien.

Fall-Kontroll-Studie:

Studientyp, bei dem Fälle einer bestimmten Krankheit mit geeigneten Kontrollen ohne diese Krankheit bezüglich der Expositionen in der Vergangenheit verglichen werden. Dient der Ermittlung von Hinweisen auf Entstehung und Ursachen einer Krankheit. Das Effektmaß ist die Odds Ratio.

Fehler 1. Art (α -Fehler):

Ein statistischer Test entdeckt einen Unterschied, obwohl er in der Realität nicht vorhanden ist (Rückweisung der Nullhypothese zu Unrecht).

Fehler 2. Art (β -Fehler):

Ein statistischer Test entdeckt keinen Unterschied, obwohl in der Realität ein Unterschied vorhanden ist (Die Nullhypothese wird zu Unrecht *nicht* zurückgewiesen.)

Fehlerquellen:

In der Epidemiologie können verschiedene Arten von Fehlern vorkommen: Eine häufige Fehlerquelle ist die Verzerrung (siehe Bias). Zufällig verteilte Fehler kommen z. B. durch die Streuung beim Messen zustande. Systematische Fehler beruhen nicht auf Zufall, sondern z. B. auf falsch eingestellten Messinstrumenten.

Fragebogen:

Ein strukturiertes Instrument zur Befragung von Probanden, das häufig in epidemiologischen Studien zur Datenerhebung benutzt wird.

Interaktion:

Das Zusammenwirken von zwei oder mehr Faktoren im Sinne der Verstärkung (Synergismus) oder Verhinderung (Antagonismus) eines Effektes.

Interventionsstudien:

Studien der experimentellen Epidemiologie, die sich durch Testung eines oder mehrerer experimenteller Faktoren (Interventionen) auszeichnen.

Inzidenz (Neuerkrankungsziffer):

Zahl neuauftretender Fälle in einer definierten Bevölkerung pro Zeiteinheit (meist pro Jahr), bezogen auf die definierte Bevölkerung (meist pro 1.000 oder pro 100.000). Im engeren Sinn: kumulative Inzidenz, auch Absolutes Risiko genannt.

Konfidenzintervall:

Aus den erhobenen Daten berechnetes Intervall, das mit einer bestimmten Wahrscheinlichkeit (in der Regel 95%) den tatsächlichen Wert einer epidemiologischen Maßzahl (wie etwa Prävalenz, Inzidenz oder Relative Risiko) beinhaltet.

Konsistenz:

Hohe Übereinstimmung bei wiederholten Messungen.

Korrelationskoeffizient:

Mathematisches Maß für den linearen Zusammenhang zweier Variablen. Der Korrelationskoeffizient wird mit dem Buchstaben r bezeichnet und kann Werte von +1 bis -1 annehmen, wobei $r = +1$ einen gänzlich positiven und $r = -1$ einen gänzlich negativen linearen Zusammenhang zweier Variablen anzeigt.

Meta-Analyse:

Quantitatives Verfahren, mit dem die Ergebnisse verschiedener Einzeluntersuchungen mit gleicher Thematik zusammengefasst werden. Ziel der Metaanalyse ist es, einen Überblick über den aktuellen Stand der Forschung zu erhalten und einen übergreifenden Trend der Ergebnisse zu identifizieren.

Monitoring:

Regelmäßige Durchführung und Analyse von Messungen zur Entdeckung von Veränderungen in der Umwelt und im Gesundheitszustand der Bevölkerung (siehe auch Surveillance = Überwachung).

Morbidität:

Maß für die Häufigkeit von Krankheit in der Bevölkerung ohne Unterscheidung zwischen Inzidenz und Prävalenz.

Mortalität

siehe Sterblichkeit

Multivariate Analysen:

Statistische Verfahren, bei denen die Abhängigkeit einer Zielerkrankung von mehreren Erklärungsvariablen simultan im Rahmen so genannter multivariater Modelle (z.B. Regressionsmodell) analysiert wird.

Nullhypothese:

Die Nullhypothese H_0 beim statistischen Test besagt, dass der in der Stichprobe gefundene Effekt auf die zufallsabhängige Streuung zurückzuführen ist und damit in der Grundgesamtheit nicht vorhanden ist.

Odds Ratio:

Der Begriff stammt aus dem Englischen. Dort spricht man bei einer Erfolgchance von einem Treffer bei zehn Versuchen nicht von einer Trefferwahrscheinlichkeit von 10%, sondern man sagt die „Odds“ für einen Treffer liegen bei 1 zu 9 oder 1/9 (im Deutschen auch „Chance“ genannt). Die Odds Ratio wird meistens im Rahmen von Fall-Kontroll-Studien berechnet und ist ein Maß für das Relative Risiko einer Erkrankung bei einer gegebenen Exposition. Es werden die „Odds“ für eine Exposition bei (erkrankten) Fällen (A/B) und (nicht erkrankten) Kontrollen (C/D) verglichen. Aus deren Quotient (Ratio) errechnet sich die Odds Ratio wie folgt:

$$\text{Odds Ratio} = (A/B) / (C/D) = (A \cdot D) / (B \cdot C)$$

Vierfeldertafel zur Berechnung der Odds Ratio in einer Fall-Kontroll-Studie

	Exposition		
	ja	nein	
Fälle	A	B	OR = $\frac{A \cdot D}{B \cdot C}$
Kontrollen	C	D	

Ökologische Studie:

Studie, bei der die Untersuchungseinheiten nicht Individuen, sondern Aggregationen (Zusammenfassungen) von Individuen sind. Gängige Aggregationen sind im geografischen Sinne administrative Einheiten (Landkreise, Länder, etc.) und im zeitlichen Sinne Jahreszusammenfassungen. Beispiel einer Ökologischen Studie ist die Assoziation von durchschnittlichem Pro-Kopf-Verbrauch von Fleisch und Sterblichkeit an Dickdarmkrebs in 23 verschiedenen Ländern.

Power:

Die statistische *Power* (dt.: Macht) bezeichnet die Wahrscheinlichkeit einer Studie, einen tatsächlich vorhandenen Zusammenhang oder Unterschied zu entdecken. Die Power einer Studie hängt von der Stichprobengröße, dem Studiendesign, der Häufigkeit der untersuchten Endpunkte (z. B. Auftreten einer Erkrankung, Tod) und der Größe des Effektes ab.

Prävalenz:

Häufigkeit von Fällen einer bestimmten Krankheit zu einem bestimmten Zeitpunkt, bezogen auf die Gesamtstichprobe oder Bevölkerung (= Punktprävalenz). Wird eine längere Zeiteinheit gewählt, so kann z. B. über eine Wochen- oder auch 12-Monatsprävalenz gesprochen werden (= Periodenprävalenz, v. a. bei rekurrierenden Erkrankungen wie Asthma).

Prävention:

Verhütung von Krankheit, häufig gleichbedeutend mit „primärer Prävention“ verwendet. Allgemein ist aber eine Unterteilung in primäre (= Erhaltung der Gesundheit und Verhinderung von Krankheit), sekundäre (= Früherkennung mit dem Ziel der Heilung) und tertiäre (= Reduktion von Behinderung) Prävention gebräuchlich.

Primäre Prävention:

Erhaltung der Gesundheit (salutogenetischer Ansatz) und Verhütung von Krankheit (pathogenetischer Ansatz) durch Stärkung gesundheitsfördernder und Beseitigung krankmachender Faktoren, Erhöhung der Resistenz von Individuen und Veränderung von Umweltfaktoren, die ursächlich oder als Überträger an der Krankheitsentstehung beteiligt sind.

Prospektive Kohortenstudie:

Studie, in der eine (Studien-) Population über einen vorgegebenen Zeitraum hinsichtlich des Eintretens interessierender Ereignisse (Erkrankung, Tod) beobachtet wird. Zu Beginn der Studie wird jeder Person ein Expositionsstatus bezüglich eines oder mehrerer interessierender Faktoren zugeordnet. Alle Studienteilnehmer müssen zu Beginn der Studie frei von den zu untersuchenden Erkrankungen sein. Prospektiv bedeutet, dass die relevanten Krankheitsereignisse sich erst in der Zukunft ereignen werden. Prospektive Kohortenstudien verkörpern eine wichtige Methode der analytischen Epidemiologie.

Querschnittsstudie:

Studie, bei der Beobachtungen über Exposition, Risikofaktoren und Krankheit in einer definierten Population zum gleichen Zeitpunkt gewonnen werden. Aufgrund des gleichzeitigen Erhebungszeitpunktes des Krankheitsstatus und der möglicherweise mit der Krankheit assoziierten Faktoren (Risikofaktoren, Exposition) eignet sich die Querschnittsstudie nur bedingt zu Kausalitätsnachweisen. Ihre Bedeutung besteht hauptsächlich in der deskriptiven Epidemiologie, besonders zur Beschreibung von Prävalenzen und zur Formulierung von Forschungshypothesen.

Randomisierung:

Zuordnung nach Zufallsprinzip (Zufallszahlentabelle; random number table) zu Untersuchungs- oder Kontrollgruppe in randomisierten klinischen Studien. Sichern die Gleichartigkeit der beiden zu vergleichenden Gruppen hinsichtlich prognoserelevanter Faktoren (so genannte Strukturgleichheit) mit Ausnahme der experimentellen Intervention.

Randomisierte klinische Studie (RCT = randomised controlled trial):

Ein sorgfältig geplantes, in Übereinstimmung mit ethischen Richtlinien durchgeführtes Experiment zur Überprüfung von Hypothesen mittels einer Zufallszuteilung der Studienteilnehmer in mindestens zwei Gruppen. Höchste Evidenzstufe zum Beleg der Wirksamkeit einer Intervention.

Registerstudie:

Studie, die auf den Daten eines oder mehrerer Register (z. B. Krebsregister, Herzinfarktregister) basiert. In einem Register werden alle Fälle einer oder mehrerer Erkrankungen oder anderer gesundheitsrelevanter Faktoren, die innerhalb einer (meist geographisch) definierten Population auftreten, erfasst. Somit können bevölkerungsbezogene Inzidenzen berechnet werden. Bei regelmäßigem Follow-up der Krankheitsfälle können auch Informationen zu Remissionen und Krankheitsrückfällen, zu Prävalenzen und zur Sterblichkeit gewonnen werden.

Relatives Risiko:

Verhältnis der Inzidenzraten (als kumulative Inzidenz oder als Inzidenzdichte) unter exponierten (I_E) gegenüber nicht exponierten Individuen (I_N).

$$RR = \frac{I_E}{I_N}$$

I_E = Inzidenz unter Exponierten
 I_N = Inzidenz unter Nicht-Exponierten

Reliabilität:

Ausmaß der Stabilität eines Resultates, wenn die zugrunde liegende Messung unter identischen Bedingungen wiederholt wird. Mangelnde Reliabilität ist die Folge von Unterschieden bei Untersuchern und / oder ihren Geräten sowie der biologischen intra-individuellen Variabilität.

Residual Confounding:

Verzerrung, die auch nach (missglückten) Versuchen, Confounding (Störfaktoren) bei der statistischen Analyse zu kontrollieren, weiter besteht. Residual Confounding entsteht, wenn unzureichende Informationen über die Störfaktoren vorhanden sind oder diese falsch kategorisiert oder missklassifiziert werden.

Risiko:

Die Wahrscheinlichkeit, dass ein Individuum innerhalb einer bestimmten Zeitperiode erkrankt oder stirbt (auch kumulative Inzidenz).

Risikofaktor:

Faktor (von Person, Umwelt oder Umstand), dessen Vorhandensein mit einer erhöhten Wahrscheinlichkeit einhergeht, eine bestimmte Krankheit oder Gesundheitsstörung zu entwickeln (z.B. Übergewicht und Diabetes). Ein Risikofaktor muss nicht unmittelbare Ursache der Gesundheitsstörung sein.

Säkularer Trend:

Veränderungen der Erkrankungs- bzw. Sterberaten über eine lange Zeitperiode (Jahre bis Jahrzehnte). Zum Beispiel Rückgang der Tuberkulose im 20. Jahrhundert oder die Zunahme der koronaren Herzkrankheit in den ersten beiden Dritteln und die Abnahme im letzten Drittel des 20. Jahrhunderts in den westlichen Industrieländern.

Sekundäre Prävention:

Früherkennung von Krankheit zu einem Zeitpunkt, da die Betroffenen noch keine Symptome wahrnehmen (präklinisches Stadium). Dazu werden Filter-Untersuchungen (Screenings) mit geeigneten Verfahren in bestimmten Zielgruppen durchgeführt (z.B. Mammographie bei 50–69-jährigen Frauen).

Sterblichkeit = Mortalität:

In einer Bevölkerung jährlich gestorbene Personen, bezogen auf die mittlere jährliche Bevölkerung in dieser Region. Meist ausgedrückt als Rohe Sterberate

$$= \frac{\text{verstorbene Einwohner in einer Region}}{\text{durchschnittliche Einwohnerzahl in derselben Region}} \times 1.000 \text{ pro Jahr.}$$

Surveillance (Überwachung):

Methoden zur Überwachung von Erkrankungen und Risikofaktoren im Gesundheitssystem durch Meldepflicht und ähnliche Aktivitäten (z.B. Sentinel-Projekte).

Survey:

Ein Survey ist eine als Querschnittsstudie durchgeführte, nicht experimentelle Erhebung zu einem ganzen Komplex von interessierenden Merkmalen. Ein Bevölkerungs-Survey kann als Interview, als schriftliche Befragung, als medizinische Untersuchung oder auch als Kombination der genannten Formen durchgeführt werden. Nationale Gesundheits-Surveys (z.B. der deutsche Bundes-Gesundheits-Survey oder der US-amerikanische National Health and Nutrition Examination Survey NHANES) sollen vielfältige Informationen über den Zustand und die Entwicklung der gesundheitlichen Lage der entsprechenden Bevölkerung liefern.

Validität:

Ausmaß der Übereinstimmung von Resultat einer Messung oder Untersuchung und dem, was sie zu messen vorgibt.