



Epidemiologie und Gesundheitsberichterstattung

Begriffe, Methoden, Beispiele
Handlungshilfe: GBE-Praxis 2

GBE-Praxis: Handlungshilfen für die Gesundheitsberichterstattung

GBE-Praxis ist als Veröffentlichungsreihe konzipiert, die methodische Hilfestellungen für die Gesundheitsberichterstattung der Landkreise und kreisfreien Städte in Bayern geben soll und in loser Reihenfolge erscheinen wird.

Das vorliegende Heft soll eine Hilfestellung bei der Anwendung epidemiologischer Methoden in der Gesundheitsberichterstattung geben: was sind die wichtigsten Maßzahlen, wie werden sie berechnet, auf welche Fehlerquellen ist zu achten, wo kann ich Rat und Hilfe bekommen?

Inhaltliche Grundlage des Kapitels 2 dieses Heftes ist das „Handbuch Epidemiologie“ von Hermann Brenner, Sebastian Schneeweis, Manfred Wildner, Uwe Siebert, Eva Grill und Karin Drooff, herausgegeben vom Bayerischen Forschungs- und Aktionsverbund Public Health e.V., München 2002, ISSN 1435-8336. Auszüge dieses Handbuchs wurden für dieses Heft mit Zustimmung des Herausgebers überarbeitet und inhaltlich an die Bedürfnisse der Gesundheitsberichterstattung in Bayern angepasst. Kapitel 3 wurde teilweise aus dem Methodenanhang des Indikatorenansatzes für die Gesundheitsberichterstattung der Länder, Bielefeld 2003, ISBN 3-88139-121-5, übernommen. Rolf Annuß vom Landesinstitut für den Öffentlichen Gesundheitsdienst Nordrhein-Westfalen danken wir für seine Unterstützung zu diesem Kapitel.

Herausgeber:

Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit
Eggenreuther Weg 43
91058 Erlangen

Telefon: 09131/764-0
Telefax: 09131/764-102

E-Mail: poststelle@lgl.bayern.de
Internet: www.lgl.bayern.de

Autorinnen und Autoren des Berichts:

Joseph Kuhn, Annette Heißenhuber, PD Dr. Manfred Wildner

Bei fachlichen Fragen wenden Sie sich bitte an:

Joseph Kuhn
Tel. 089/31560-302
joseph.kuhn@lgl.bayern.de

Stand:

September 2004

Diese Druckschrift wird kostenlos im Rahmen der Öffentlichkeitsarbeit der Bayerischen Staatsregierung herausgegeben. Sie darf weder von den Parteien noch von Wahlwerbern oder Wahlhelfern im Zeitraum von fünf Monaten vor einer Wahl zum Zweck der Wahlwerbung verwendet werden. Dies gilt für Landtags-, Bundestags-, Kommunal- und Europawahlen. Missbräuchlich ist während dieser Zeit insbesondere die Verteilung auf Wahlveranstaltungen, an Informationsständen der Parteien sowie das Einlegen, Aufdrucken und Aufkleben parteipolitischer Informationen oder Werbemittel. Untersagt ist gleichfalls die Weitergabe an Dritte zum Zweck der Wahlwerbung. Auch ohne zeitlichen Bezug zu einer bevorstehenden Wahl darf die Druckschrift nicht in einer Weise verwendet werden, die als Parteinahme der Staatsregierung zugunsten einzelner politischer Gruppen verstanden werden könnte. Den Parteien ist es gestattet, die Druckschrift zur Unterrichtung ihrer eigenen Mitglieder zu verwenden. Bei publizistischer Verwertung – auch von Teilen – Angabe der Quelle und Übersendung eines Belegexemplars erbeten. Das Werk ist urheberrechtlich geschützt. Alle Rechte sind vorbehalten. Die Publikation wird kostenlos abgegeben, jede entgeltliche Weitergabe ist untersagt. Diese Druckschrift wurde mit großer Sorgfalt zusammengestellt. Eine Gewähr für die Richtigkeit und Vollständigkeit kann dennoch nicht übernommen werden.

Inhalt

1. Einführung: Epidemiologie und Gesundheitsberichterstattung	4
2. Grundbegriffe und Maßzahlen der Epidemiologie	6
2.1. Kausalität und Assoziation	6
2.2. Prävalenz und Inzidenz	8
2.3. Effektmaße	12
3. Statistische Methoden in der Epidemiologie	18
3.1. Altersstandardisierung	18
3.2. Signifikanzprüfung	22
3.3. Weitere Verfahren zur Darstellung der Sterblichkeit	23
4. Epidemiologische Studien	26
4.1. Studientypen: ein kurzer Überblick	26
4.2. Die Qualität von Studien	27
4.3. Evidenzklassen	29
4.4. Objektivität, Reliabilität, Sensitivität und Spezifität	30
5. Literatur	34

1. Einführung: Epidemiologie und Gesundheitsberichterstattung

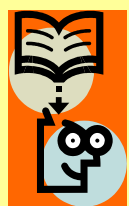
Das Wort Epidemiologie bedeutet übersetzt die Lehre davon, „was auf dem Volk liegt“. In dieser Begriffsbildung klingt noch die Herkunft der Epidemiologie aus der Seuchenbekämpfung nach. Heute geht es ganz allgemein um eine Methodenlehre zur Darstellung der Häufigkeit und Verteilung von Risikofaktoren, Erkrankungen und Befindlichkeitsstörungen in der Bevölkerung. So haben es z.B. auch die amerikanischen Epidemiologen MACMAHON und PUGH (1970) formuliert:

Epidemiologie ist „... the study of the distribution and determinants of disease frequency in human populations“

Definition

Die *deskriptive Epidemiologie* beschreibt das Auftreten von Krankheiten in einer bestimmten Bevölkerung im Zusammenhang mit soziodemografischen Variablen wie Alter, Geschlecht, Beruf oder Umwelt. Die deskriptive Epidemiologie ist die Voraussetzung der analytischen Epidemiologie.

Die *analytische Epidemiologie* geht den Hintergründen der Erkrankung nach. Im Zentrum steht dabei die Frage nach Ursache und Wirkung. Aufbauend auf der Analyse von Daten, die die deskriptive Epidemiologie geliefert hat, werden Hypothesen entwickelt und geprüft.



Ein Beispiel: In einem Landkreis werden im Rahmen der Gesundheitsberichterstattung routinemäßig die Krebserkrankungen bei Kindern beobachtet. Das Beschreiben der Häufigkeit der Erkrankungen, der Unterschiede nach Jungen und Mädchen, des zeitlichen Verlaufs und des regionalen Vergleichs mit den Nachbarlandkreisen stellt eine Form der deskriptiven Epidemiologie dar. Dazu braucht man vor allem die Absolutzahlen der Erkrankungen, die dazugehörige Bevölkerungszahl und soziodemografische Merkmale zur Differenzierung der Fälle. Gesetzt den Fall, das Gesundheitsamt beobachtet nun eine ungewöhnliche Häufung von Leukämiefällen und möchte wissen, ob diese Häufung statistisch signifikant ist und ob es Zusammenhänge mit einer Umweltexposition gibt, so kommen zur Beantwortung dieser Fragen Methoden der analytischen Epidemiologie zur Anwendung.

Die Epidemiologie ist eine wichtige Grundlage der Gesundheitsberichterstattung, Gesundheitsberichterstattung ist aber nicht identisch mit Epidemiologie. Während die Epidemiologie eine wissenschaftliche Methode darstellt und zunächst einmal im wissenschaftlichen Kontext beheimatet ist, ist die Gesundheitsberichterstattung eher ein gesundheitspolitischer Kommunikationsprozess, der versucht, eine Brücke zwischen Wissenschaft, Politik und (Fach-)Öffentlichkeit zu schaffen. Wo die Epidemiologie differenziert und analysiert, da aggregiert und systematisiert die Gesundheitsberichterstattung (siehe dazu auch GBE-Praxis Heft 1).

**Epidemiologie:
eine Grundlage
der GBE**

Zur Verdeutlichung der unterschiedlichen Akzentuierung von Epidemiologie und Gesundheitsberichterstattung kann folgende Tabelle dienen (vgl. auch Kuhn/Wildner 2003).

Gesundheitsberichterstattung und Epidemiologie		
	Epidemiologie	GBE
Ziel	Neues entdecken	Verhältnisse beobachten, Transparenz herstellen, Diskussion anregen
Datenauswahl	Wissenschaftliche Fragen, Forschungsfinanzierung	Politische Schwerpunktsetzung
Regionalbezug	Gelegentlich	Immer
Adressaten	Wissenschaft, Auftraggeber	Politik, Bürger, Verwaltung
Sprache	Wissenschaftlich exakt, mit ausführlichen methodischen Erläuterungen	Allgemeinverständlich
Produzenten	Forschungsgruppen	Verwaltung

**Unterschiede
zwischen GBE und
Epidemiologie**

Die Epidemiologie kann man als einen Zweig der Methodenlehre der gesundheitswissenschaftlichen Forschung verstehen. Die verschiedenen, damit einhergehenden grundsätzlichen Fragestellungen (Theorie-Praxis-Zusammenhang, Begriffsbildung, Hypothesengenerierung, Messtheorie, Datenerhebung, Datenauswertung etc.) können hier nicht im Einzelnen behandelt werden, dazu sei auf die weiterführende Literatur verwiesen.

Dieses Heft soll lediglich einen kurzen Überblick über die wichtigsten epidemiologischen Kennziffern geben und helfen, vorhandene epidemiologische Studien besser beurteilen und für die Gesundheitsberichterstattung nutzen zu können. Für die Planung eigener Studien, vor allem wenn es um den Nachweis kleiner Effekte oder die Analyse kleiner Fallzahlen geht, empfiehlt es sich in der Regel, externe Beratung zu suchen. Das Bayerische Landesamt für Gesundheit und Lebensmittelsicherheit unterstützt die Landkreise und kreisfreien Städte dabei, ggf. auch durch die Vermittlung von Fachleuten.

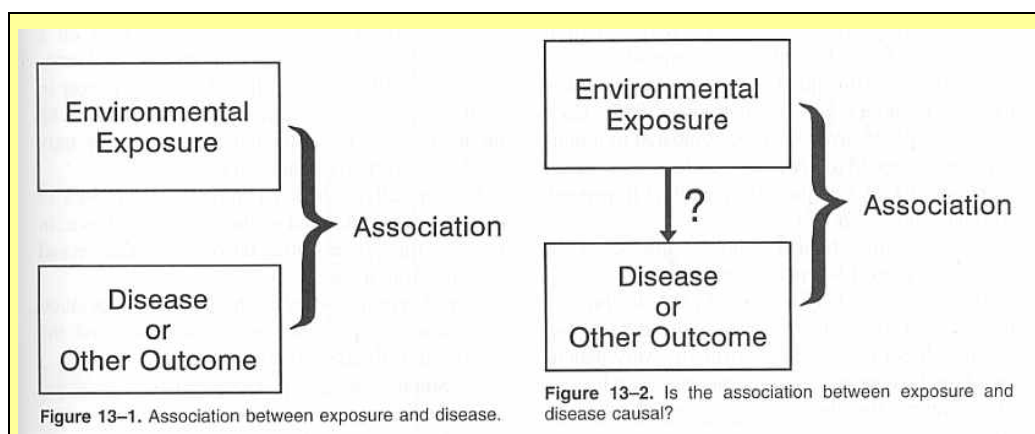
2. Grundbegriffe und Maßzahlen der Epidemiologie

2.1. Kausalität und Assoziation

Die Epidemiologie will, wie erwähnt, zur Aufklärung von Krankheitsursachen beitragen und bedient sich dazu verschiedener statistischer Methoden. Eine festgestellte statistische Beziehung zwischen zwei Merkmalen muss jedoch noch keinen kausalen, ursächlichen Zusammenhang bedeuten. Man spricht daher, wenn eine Beziehung statistisch besteht, aber (noch) nicht kausal interpretiert werden kann, von einer „**Assoziation**“ von Merkmalen. So lässt sich in den Schulinganguntersuchungen beispielsweise eine statistisch signifikante Korrelation zwischen kindlichem Übergewicht und Sprachstörungen feststellen. Medizinisch hat das eine mit dem anderen aber in der Regel nichts zu tun. Der statistische Zusammenhang kommt dadurch zustande, dass auf beide Merkmale gemeinsame Einflussfaktoren wirken, z.B. die soziale Lage der Kinder. Ein berühmtes Beispiel zur Verdeutlichung des Unterschieds zwischen einer Assoziation von Merkmalen und kausalen Zusammenhängen ist die statistische Korrelation des Rückgangs von Störchen und Geburten in den letzten Jahrzehnten. Beides hat kausal bekanntlich nichts miteinander zu tun.

**Zusammenhangs-
analyse**

Eine Assoziation von Merkmalen stellt also zunächst nur einen Hinweis auf einen möglichen kausalen Zusammenhang dar, den es zu bestätigen oder zu widerlegen gilt:



**Kausalität und
Assoziation**

Quelle: GORDIS 1996: 167

Um den „kausalen Gehalt“ von epidemiologischen Beobachtungen abzuschätzen, kann man sich an den Kausalitätskriterien nach AUSTIN BRADFORD HILL (1965) orientieren:

Kausalitätskriterien nach BRADFORD HILL

1. Stärke der Beziehung	Schwache Assoziation: Zufall?
2. Konsistenz der Beziehung	Ist die Beziehung wiederzufinden in anderen Studien, Populationen und unter verschiedenen Bedingungen?
3. Spezifität des Effekts	Ist der Effekt klar abgrenzbar? Viele Effekte im Zusammenhang mit der Ursache?
4. Zeitliche Sequenz	Ist die Reihenfolge Exposition-Krankheit gegeben?
5. Dosis-Wirkungs-Beziehung	Gibt es eine höhere Wirkung bei höherer Dosis?
6. Biologische Plausibilität	Ist die Kausalbeziehung biologisch plausibel?
7. Kohärenz	Steht ein angenommener Zusammenhang in Widerspruch zum aktuellen Wissensstand von Ätiologie und Pathogenese der Erkrankung?
8. Experimentelle Evidenz	Eine neue Hypothese sollte prüfbar sein. Ist eine solche Prüfung durchgeführt worden?
9. Analogie mit ähnlichen Kausalzusammenhängen	Gibt es ähnliche Kausalzusammenhänge bei ähnlichen Sachverhalten?

Kausalitätskriterien

Jedoch müssen solche „Faustregeln“ vorsichtig angewendet werden. Insbesondere die Stärke einer statistischen Korrelation sagt nicht immer etwas über den kausalen Charakter eines Zusammenhangs aus, sondern kann z.B. auch ein Hinweis auf einen „Confounder“, eine den beiden korrelierenden Merkmalen gemeinsame Ursache, sein.

Ein weiteres Problem besteht darin, dass beim Entstehen einer Erkrankung meist mehrere Einflussfaktoren zusammen wirken. Dieses Zusammenspiel folgt nur selten einer einfachen linearen Logik („A verursacht B, B verursacht C“), sondern ist oft komplexerer Natur. ROTHMANN 1976 hat dazu eine hilfreiche Einteilung vorgeschlagen. Er unterscheidet drei Typen von Kausalbeziehungen:

- **Hinreichende Krankheitsursachen („sufficient cause“)**
Zusammentreffen einer Reihe von Bedingungen bzw. Ereignissen, die unweigerlich zum Krankheitsausbruch führen.
- **Teilursachen („component cause“)**
Bedingung, die Bestandteil einer hinreichenden Krankheitsursache ist, für sich genommen aber nicht zur Verursachung der Krankheit führt.
- **Notwendige Ursachen („necessary cause“)**
Bedingung, die Bestandteil aller hinreichenden Krankheitsursachen ist, d.h. ohne die die Entstehung der Krankheit nicht möglich ist.

Ursachenbündel



Beispiel:

Ursachengeflecht einer Tuberkuloseerkrankung

1. Notwendige Ursache: Infektion mit Mykobakterium
2. Teilursachen: Lebensumstände, Ernährungszustand, Lebensstil
3. Hinreichende Ursache: Zusammentreffen von 1 und 2

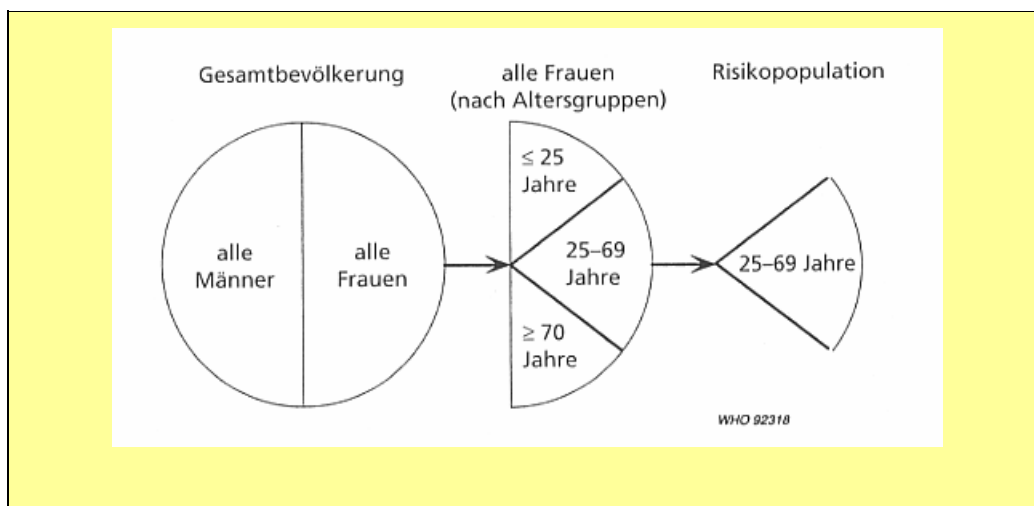
2.2. Prävalenz und Inzidenz

Eine wesentliche Aufgabe der Epidemiologie ist die *quantitative Erfassung* von Ereignissen und Effekten und ihre Zusammenfassung in aussagekräftige Maßzahlen.

Eine erste wichtige Voraussetzung für die Berechnung aller epidemiologischen Maßzahlen ist dabei die möglichst genaue Definition von Fällen und Bezugspopulation. Dabei sind die Fälle stets eine Teilmenge der Bezugspopulation. Die Bezugspopulation wiederum soll so gewählt werden, dass sie möglichst nur Personen umfasst, die überhaupt zu Fällen werden können, z.B. Personen, die für die untersuchte Krankheit potentiell anfällig sind.

Beispielsweise würden in einer Studie über das Auftreten des Zervixkarzinoms Männer nicht miteingeschlossen:

Risikopopulation in einer Studie zum Auftreten des Zervixkarzinoms



Quelle: BEAGLEHOLE, BONIITA, KJELLSTRÖM 1997: 30.

**Fälle und
Bezugspopulation**

Die Beschreibung der **Häufigkeiten** von Erkrankungen kann auf zwei Aspekte abheben: den *Bestand* und den *Neuzugang* an Fällen. Erstes ist die Prävalenz, letzteres die Inzidenz der Erkrankung.

Prävalenz

Die **Prävalenz** gibt den Anteil Erkrankter zu einem bestimmten Zeitpunkt (**Punktprävalenz**) oder innerhalb eines gewissen Zeitraums (**Periodenprävalenz**) an. Sie ist also eine Verhältniszahl.

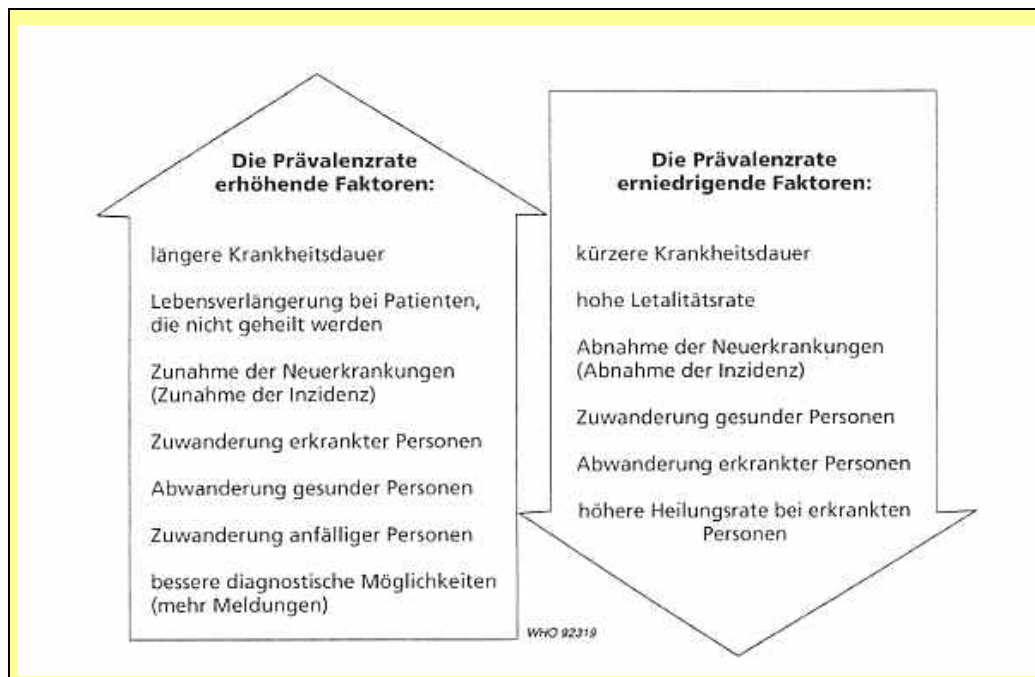
Prävalenz

$$\text{Prävalenz } P = \frac{\text{Anzahl der Fälle}}{\text{Gesamtpopulation}}$$

Die Prävalenz unterliegt dabei dem Einfluss verschiedener Faktoren:

- der *Schwere der Krankheit* (wenn zahlreiche Erkrankte sterben, nimmt die Prävalenz ab);
- der *Krankheitsdauer* (bei kurzer Krankheitsdauer ist die Prävalenz niedriger als bei langer Krankheitsdauer);
- der *Zahl der Neuerkrankungen* (wenn viele Menschen erkranken, ist die Prävalenz höher als bei wenigen Erkrankten).

Einflussfaktoren



Quelle: BEAGLE, BONITA, KJELLSTRÖM 1997: 32

Inzidenz

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

Zur Messung der Inzidenz gibt es zwei grundlegende Inzidenzmaße. Das erste Maß ist die **kumulative Inzidenz** oder das **absolute Risiko**, die für einen definierten Zeitabschnitt und eine bestimmte Anfangsbevölkerung (**Kohorte**) berechnet wird.

$$\text{Kumulative Inzidenz } I_c = \frac{\text{Anzahl der Neuerkrankungen}}{\text{Bezugsbevölkerung}}$$

Die kumulative Inzidenz wird als Anteil ausgedrückt und ist dimensionslos. Es muss dabei stets eine Angabe über den Bezugszeitraum gemacht werden. Beispiel: Das Risiko (die I_c), in den *nächsten 10 Jahren* zu erkranken, beträgt 30%. Die kumulative Inzidenz liegt zwischen 0 und 1 (0 bzw. 100%).

In der Gesundheitsberichterstattung genügt für die Abschätzung der Inzidenz von Erkrankungen meist diese Maßzahl, wobei in vielen Fällen als Anfangsbevölkerung die Gesamtbevölkerung genommen wird. Das ist immer dann zulässig, wenn die darin natürlich auch enthaltenen Kranken oder die Zu- und Abwanderungen statistisch nicht ins Gewicht fallen. Bei vielen epidemiologischen Studien und kleineren Fallzahlen ist es dagegen notwendig, die Studienbevölkerung zu Beginn genau festzulegen. Man spricht dann vom Design einer sogenannten **geschlossenen Kohorte** (*closed/fixed cohort*).

Es ist aber nicht immer gewährleistet, dass eine Kohorte in einer Studie gleich bleibt. Personen können wegziehen oder auch sterben. Um trotzdem zu realistischen Maßzahlen zu gelangen, berechnet man die **Inzidenzdichte**. Sie bezieht sich nicht mehr auf die Anfangsbevölkerung, sondern auf die innerhalb dieser Bevölkerung im Studienverlauf beobachtete Personenzeit. Ein Beispiel: Zum Vergleich des Risikos von Arbeitsunfällen zwischen verschiedenen Branchen bezieht man häufig die in einer Periode gemeldeten Arbeitsunfälle auf die Personenzeit, hier die geleisteten Arbeitsstunden. Die Inzidenzdichte zählt Ereignisse pro Personenzeit, ist also eine Rate. Ihre Einheit sind zumeist „Neuerkrankungen pro Personenjahre“. Studienkonzept ist die sogenannte **offene oder dynamische Kohorte** (*open/dynamic cohort*).

Kumulative Inzidenz

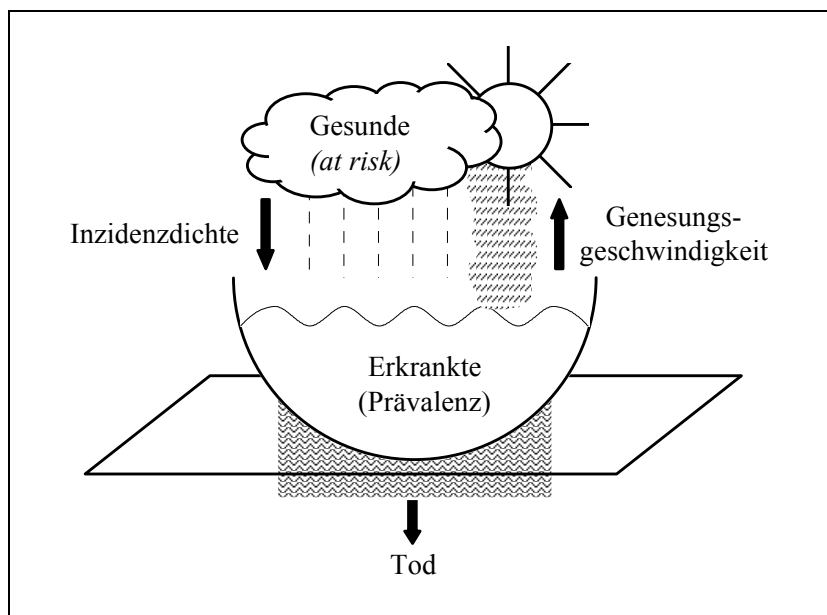
Inzidenzdichte

$$\text{Inzidenzdichte } I_d = \frac{\text{Anzahl der Neuerkrankungen}}{\text{Personenzeit}}$$

(Personenzeit: Population unter Risiko mal Zeitraum)

Zusammenhang von Prävalenz und Inzidenz

Die vorgestellten Maßzahlen Prävalenz und Inzidenz (kumulative Inzidenz und Inzidenzdichte) können unter bestimmten Voraussetzungen und bei weiteren gegebenen Informationen ineinander umgerechnet werden. Die folgende Abbildung veranschaulicht den Zusammenhang zwischen Prävalenz, Inzidenzdichte und mittlerer Krankheitsdauer anhand des natürlichen Wasserkreislaufes.



Der „Prävalenzpool“

Stellt man sich die Menge der Erkrankten als Wassermenge in einem See bzw. Behälter vor, der sowohl Zu- als auch Abfluss besitzt, dann steigt die *Prävalenz* (Wassermenge) bei größerer *Inzidenzdichte* (Zuflussgeschwindigkeit) und *längerer Krankheitsdauer* (Verweildauer). Sowohl schnelle Genesung als auch schneller Todeseintritt (Abfluss) führen zu einer kurzen Krankheitsdauer und vermindern die Prävalenz.

Der beschriebene Zusammenhang nimmt eine einfache Form an, wenn man von einem „Steady State“ ausgeht, d.h. von der Annahme, dass sich zum Zeitpunkt der Untersuchung ein Gleichgewichtszustand eingestellt hat, bei dem die Zuwanderungen zum *Prävalenzpool* mit den Abwanderungen (Genesungen) gleich sind.

Prävalenz $P \approx I_d \times D$; diese Näherung gilt für niedrige Prävalenzen, D ist die mittlere Dauer der Erkrankung

$$\begin{aligned} \text{Inzidenzdichte } I_d &= \frac{\text{Neuerkrankungen}}{(\text{Population unter Risiko}) \times \text{Zeitraum}} \\ &\approx \frac{I_c}{\text{Beobachtungszeitraum}} \end{aligned}$$

Des Weiteren kann man bei akuten und nicht lange andauernden Erkrankungen die Prävalenz als Schätzer für die Inzidenz benutzen, z.B. bei Influenza-Infektionen.

**Prävalenz als
Schätzer für
die Inzidenz**

2.3. Effektmaße

Relatives Risiko

Mit epidemiologischen Methoden lassen sich auch Hinweise darauf gewinnen, welche Auswirkungen eine Exposition hat. Man spricht hier vom **Effekt**, der Auswirkung einer bestimmten Exposition im Vergleich zur Nichtexposition. Dabei gibt es zwei grundlegende Effektmaße: Verhältnismaße (Raten) und Differenzmaße. Im Folgenden werden zunächst die einzelnen Maße in ihrer inhaltlichen Bedeutung kurz erläutert. Wie man die Maße aus den Ausgangsdaten konkret berechnet, wird etwas später anhand einer Vierfeldertafel gezeigt.

Risiko

Unter den Verhältnismaßen am häufigsten ist das **relative Risiko**. Darunter versteht man das Verhältnis der kumulativen Inzidenz von exponierten und nicht exponierten Personen. Das Verhältnis der jeweiligen Inzidenzdichten wird als **relative Rate** bezeichnet. Beide Größen werden als **RR** abgekürzt, was etwas verwirrend sein kann. Beide Maße geben an, um wie viel höher (oder niedriger) das Risiko zu erkranken für eine exponierte Person im Vergleich zu einer nicht-exponierten Person ist.

**Relatives Risiko,
Relative Rate**

In Formeln tragen die Nichtexponierten üblicherweise das Suffix „0“, die Exponierten das Suffix „1“ bzw. bei mehreren Expositionsklassen eine andere Zahl. Die Formeln sehen dann wie folgt aus:

$$\text{Relatives Risiko RR} = \frac{I_{c_1}}{I_{c_0}}$$

$$\text{Relative Rate RR} = \frac{I_{d_1}}{I_{d_0}}$$

Ist das Relative Risiko größer als 1, dann ist dies ein Hinweis auf einen möglichen kausalen Zusammenhang zwischen Exposition und Erkrankung. Ist das Relative Risiko kleiner als 1, dann könnte die Exposition einen protektiven Einfluss haben, also das Erkrankungsrisiko vermindern. Bei einem Relativen Risiko gleich 1 gibt es statistisch keine Assoziation zwischen Exposition und Erkrankung.



Beispiel: Ein Landkreis macht eine Studie zu Verkehrsunfällen bei jungen Männern. Von 1.000 Befragten der relevanten Altersgruppe gaben 50 an, im letzten Jahr einen Unfall gehabt zu haben. Die kumulative Inzidenz beträgt dann $50/1000$, also 0,05, also 5%. Von den 1.000 Befragten haben 120 eine nicht korrigierte Kurzsichtigkeit. Aus dieser Gruppe hatten 24 einen Unfall. I_{c_1} beträgt

dann $24/120$, also 0,2 und I_{c_0} beträgt $26/880$, also ca. 0,03.

Das Relative Risiko derjenigen mit Sehfehler gegenüber denen ohne Sehfehler errechnet sich folglich zu $0,2/0,03 = 6,67$.

Odds Ratio

Eine weitere häufige Maßzahl ist die **Odds Ratio**, die als **OR** abgekürzt wird und oft Schwierigkeiten im Verständnis macht. Die Odds Ratio benutzt man als Näherungswert für das Relative Risiko, wenn man dieses nicht direkt berechnen kann, also wenn das Studiendesign keine Daten zur Inzidenz liefert. Dieses Problem stellt sich regelmäßig in Fall-Kontroll-Studien. Die Odds gibt das Verhältnis der Häufigkeit des Eintritts eines Ereignisses zur Häufigkeit des Nichteintritts entweder unter Exponierten oder unter Kranken an.

Odds Ratio



Beispiel: In einem Landkreis sind von 100 Einschulungskindern aus sozial benachteiligten Familien (Exponierte) 15 übergewichtig, 85 nicht. Die Odds beträgt $15/85 = 0,18$. Unter 100 Kindern aus nicht sozial benachteiligten Familien sind nur 10 übergewichtig. Die Odds beträgt hier $10/90 = 0,11$. Die Odds Ratio beträgt damit $0,18/0,11 = 1,63$.

Bei seltenen Erkrankungen sind das Risiko und die Odds zahlenmäßig sehr ähnlich, so dass dann das **RR** durch die **OR** abgeschätzt werden kann. Es gilt dann: $OR \approx RR$. Die Zahlenwerte der Odds Ratio sind wie die des Relativen Risikos zu bewerten; ein OR größer als 1 deutet einen krankmachenden Effekt der Exposition hin, ein OR kleiner als 1 auf einen protektiven Effekt.

Risikodifferenz

Eine weitere Möglichkeit, Effekte der Exposition gegenüber Nichtexposition zu quantifizieren, ist die Bildung von Differenzmaßen. Hier wird wieder zwischen **Risikodifferenz** und **Ratendifferenz** unterschieden. Leider werden diese beiden Größen ebenfalls identisch als **RD** abgekürzt. Das überschüssige, „exzessive“ Risiko unter den Exponierten wird, falls es keine andere Erklärung gibt, der Exposition zugeschrieben. Die Risikodifferenz wird daher auch bezogen auf die Exposition **attributables Risiko**, also der Exposition **zuschreibbares Risiko**, genannt.

Risikodifferenz (Attributables Risiko) $RD = I_{c_1} - I_{c_0}$

Ratendifferenz (Attributable Rate) $RD = I_{d_1} - I_{d_0}$

Das attributable Risiko ist vor allem auch für die Planung von Präventionsmaßnahmen von Bedeutung, weil es angibt, in welchem Umfang die Inzidenz einer Erkrankung gesenkt werden kann, wenn die Exposition nicht mehr vorliegt.

Attributable Risiko Proportion

Eine weitere Form der Angabe von zuschreibbaren Risiken ist die **attributable Risiko Proportion** (Synonyme: Etiologic fraction (EF), attributable risk percent (ARP) oder attributables Risiko unter Exponierten (ARE)). Sie ist ein Maß für den Anteil der Krankheit, der bei den Exponierten auf die Exposition zurückzuführen ist, z.B. den Anteil an Lungenkrebsfällen unter Rauchern, der durch das Rauchen verursacht wird. Man bezieht dazu die Risikodifferenz auf die Inzidenz der Exponierten und kann dann nach einer Formelumstellung die attributable Risiko Proportion recht einfach aus dem Relativen Risiko berechnen:

$$ARP_e = \frac{I_{c_1} - I_{c_0}}{I_{c_1}} = \frac{RR - 1}{RR}$$

Voraussetzung für die Anwendung dieser Formel ist allerdings, dass das Relative Risiko $RR > 1$ ist. Für den Fall, dass das $RR < 1$ ist, berechnet man die „Preventive Fraction“ (siehe unten).

**Risikodifferenz,
attributables
Risiko**

**Attributable
Risiko Proportion**



Beispiel: In der Studie zu den Verkehrsunfällen lag das Relative Risiko der Kurzsichtigen bei 6,67. Die attributable Risiko Proportion beträgt demnach $(6,67 - 1)/6,67 = 0,85$.

Demnach wären 85 % der Unfälle der Kurzsichtigen auf die Kurzsichtigkeit zurückzuführen, 15 % dagegen auf Einflussfaktoren, denen auch die Nicht-Kurzsichtigen ausgesetzt sind.

Attributable Risiko Proportion in der Gesamtpopulation

Die nächste interessante Frage ist, welchen Anteil die expositionsbedingten Erkrankungen an den Erkrankungen insgesamt haben, also z.B. wie viel von allen Lungenkrebsfällen durch das Rauchen bedingt sind. Dazu benötigt man neben der attributablen Risiko Proportion noch die Prävalenz der Exposition unter allen Erkrankten, d.h. den Anteil der Exponierten unter allen Kranken in der Bevölkerung (p_e). Die bevölkerungsbezogene zuschreibbare Proportion ist ebenfalls für die Präventionsplanung von Interesse, weil sie abschätzt, was ein Präventionsprogramm zur Reduzierung der Exposition erreichen kann - in welchem Umfang z.B. ein Anti-Raucher-Programm maximal die Lungenkrebsfälle verringern kann.

$$ARP_g = p_e \times ARP_e = p_e \times \frac{RR - 1}{RR}$$

„Preventive fraction“ für protektive Faktoren

Wenn eine Exposition protektiv wirkt, also RR bzw. OR < 1 sind, spricht man von der präventablen Fraktion. Sie berechnet sich als

Preventive fraction

$$PF_e = \frac{I_{c_0} - I_{c_1}}{I_{c_0}} = 1 - RR \text{ (bzw. } 1 - OR)$$

$$PF_e = \frac{\text{durch Prävention (Exp.) verhinderte Fälle}}{\text{ohne Prävention (Exp.) zu erwartende Fälle}}$$



Vierfeldertafel zur Berechnung der Effektmaße

Für die Berechnung dieser Maße stellt man die Ausgangsdaten am Besten in einer Vierfeldertafel dar. Damit lassen sich sowohl geschlossene Kohortenstudien als auch offene Kohortenstudien mit Personenzeit übersichtlich darstellen:

	Erkrankt	Nicht erkrankt	Gesamt	
Exponiert	a	b	a + b	Berechnungshilfe
Nicht exponiert	c	d	c + d	

Dann berechnen sich das Relative Risiko, die Risikodifferenz und die Odds Ratio wie folgt:

$$\text{Relatives Risiko RR} = \frac{I_{c_1}}{I_{c_0}} = \frac{\left(\frac{a}{a+b}\right)}{\left(\frac{c}{c+d}\right)}$$

Relatives Risiko

$$\text{Risikodifferenz RD} = I_{c_1} - I_{c_0} = \frac{a}{a+b} - \frac{c}{c+d}$$

Risikodifferenz

$$\text{Odds Ratio OR} = \frac{\text{Verhältnis von Erkrankten zu Nichterkrankten unter Exposition}}{\text{Verhältnis von Erkrankten zu Nichterkrankten ohne Exposition}}$$

$$= \frac{\left(\frac{a}{b}\right)}{\left(\frac{c}{d}\right)} = \frac{a \times d}{b \times c}$$

Odds Ratio

Die Berechnung der Maßzahlen aus offenen Kohorten, bei denen mit Personenzeit gerechnet wird, ist ebenfalls über eine Vierfeldertafel möglich. Die Vierfeldertafel sieht dabei allerdings etwas anders aus:

	Erkrankt	Personenzeit	
Exponiert	a	n ₁	... und für offene Kohorten
Nicht exponiert	c	n ₀	

Und es berechnen sich:

$$\text{Relative Rate RR} = \frac{I_{d_1}}{I_{d_0}} = \frac{\left(\frac{a}{n_1}\right)}{\left(\frac{c}{n_0}\right)}$$

$$\text{Ratendifferenz RD} = I_{d_1} - I_{d_0} = \frac{a}{n_1} - \frac{c}{n_0}$$



Welches Effektmaß bei welcher Studie?

Die Anwendung der verschiedenen Effektmaße ergibt sich zumeist schon aus dem gewählten bzw. vorgegebenen Studiendesign. Sie kann sich aber auch aus der beabsichtigten Analysestrategie ergeben. Als Faustregeln gelten:

- RR und RD bei Kohortenstudien,
- OR bei Fall-Kontroll-Studien und bei logistischen Regressionsverfahren,
- Attributable Proportionen bei bevölkerungsbezogenen („repräsentativen“) Studien.

Die Effektmaße RR, RD und OR werden berechnet, wo Entscheidungen wie „exponiert/nicht exponiert“ und „krank/nicht krank getroffen werden können. Einflussfaktoren und Gesundheitszustand sind hier dichotom verteilt. Bei skalierten Größen (z.B. Herzfrequenz und Puls) kommen statistisch anspruchsvollere Verfahren der Quantifizierung von Zusammenhängen zur Anwendung (Korrelationskoeffizienten, Regressionsanalysen etc.), dazu sei auf die einschlägige Statistik-Literatur verwiesen.

3. Statistische Methoden in der Gesundheitsberichterstattung

Im Folgenden werden einige grundlegende statistische Verfahren für die Gesundheitsberichterstattung dargestellt. Die Darstellung ist weitgehend aus dem Anhang zum Indikatorenset der Gesundheitsberichterstattung der Länder übernommen (AOLG 2003). Als Hilfsmittel für die Berechnung der vorgestellten Maße können Excel-Arbeitsblätter mit den schon eingefügten Formeln zur Verfügung gestellt werden (siehe unten).

3.1. Altersstandardisierung

In der Gesundheitsberichterstattung ist es oft notwendig, Daten über eine längere Zeitperiode hinweg oder zwischen verschiedenen Regionen zu vergleichen. So lassen sich beispielsweise Fragen nach einem Anstieg bestimmter Erkrankungen innerhalb der letzten Jahre oder nach Unterschieden der Sterblichkeitsziffern zwischen Landkreisen untersuchen. Ein erster Analyseschritt ist dabei die Berechnung der sogenannten „rohen“ Krankheits- oder Sterbeziffern, die die Anzahl der Sterbefälle (bzw. Erkrankungen) je 100.000 Einwohner der mittleren Jahresbevölkerung angeben:

$$\text{Rohe Sterberate MR} = \frac{\text{Anzahl Verstorbene}}{\text{Durchschnittliche Bevölkerung}} \times 100.000$$

Die durchschnittliche Bevölkerung errechnet sich aus der Summe der Jahresanfangs- und Jahresendbevölkerung, dividiert durch 2. Epidemiologisch gesehen ist die rohe Sterberate eine Inzidenz. Sie ist abhängig vom Altersaufbau und der Geschlechterverteilung in der untersuchten Population. Beides hat einen starken Einfluss auf die Mortalität. Für den Vergleich des Sterbegeschehens verschiedener Populationen ist es daher sinnvoll, den Einfluss dieser Faktoren zu kontrollieren, damit andere Effekte nicht völlig überlagert werden. Hinsichtlich der Geschlechterverteilung ist das nicht schwer: man berechnet einfach die Sterberaten für beide Geschlechter getrennt. Der Einfluss des Altersaufbaus muss dagegen durch geeignete Standardisierungsverfahren bereinigt werden. Dabei werden zwei Methoden unterschieden: die direkte und die indirekte Altersstandardisierung.

**Einfluss von
Geschlecht
und Alter**

Direkte Altersstandardisierung

Das Verfahren der direkten Altersstandardisierung unterlegt der zu untersuchenden Population den Altersaufbau einer Standardbevölkerung. Die so gewonnenen direkt altersstandardisierten Mortalitätsraten entsprechen den im englischen Sprachraum verwendeten *standard death rates*, *SDR*. Für die Berechnung wird meist eine Bevölkerungsunterteilung nach 5-Jahres-Altersgruppen verwendet. Die Mortalitätsraten der einzelnen Altersgruppen in der beobachteten Bevölkerung werden hierbei mit den Bevölkerungsanteilen der Standardpopulation gewichtet entsprechend der Formel:

$$\text{Altersstandardisierte Sterberate } MR_{st} = \frac{\sum (N_i \cdot mr_i)}{\sum N_i}$$

wobei N_i die Zahl der Personen in der Altersgruppe i der Standardbevölkerung ist, mr_i für die altersspezifischen Mortalitätsraten je 100.000 der untersuchten Bevölkerung steht und \sum die Summe über alle Altersgruppen anzeigt.

Es wird also die Sterberate jeder Altersgruppe mit der Zahl der Personen in der entsprechenden Gruppe der Standardpopulation multipliziert (gewichtet); die Produkte werden summiert und durch die Gesamtpersonenzahl der Standardpopulation dividiert. Die Sterberaten werden demnach so berechnet, als wäre die Altersverteilung identisch mit derjenigen der Standardbevölkerung.



Der Effekt einer Altersstandardisierung lässt sich gut anhand der Entwicklung der Sterblichkeit von Frauen innerhalb der letzten 20 Jahre veranschaulichen. Während die rohe Mortalitätsrate der Frauen in dieser Zeit praktisch unverändert blieb, zeigen die altersstandardisierten Raten einen Rückgang der Mortalität um 30 %. Somit ist der Verlauf der rohen Sterblichkeitsrate in erster Linie durch Veränderungen der Altersstruktur zu erklären, d.h. den gestiegenen Anteil älterer Frauen mit höherer Sterblichkeit unter der Gesamtheit aller Frauen. Dies hat den gesundheitspolitisch bedeutsamen Rückgang der Sterblichkeit in allen einzelnen Altersklassen überlagert.

In der Gesundheitsberichterstattung wird als Standardbevölkerung die „alte Europa-Standardbevölkerung“ benutzt. Da diese alte Europa-Standardbevölkerung von 1966 im Altersaufbau inzwischen stark vom tatsächlichen Altersaufbau der europäischen Bevölkerung abweicht, wurde 1990 eine „neue“ Europa-Standardbevölkerung vorgeschlagen. Diese „Neue Europäische Standardbevölkerung“ hat sich in der Gesundheitsberichterstattung sowohl international als auch innerhalb Deutschlands aber nicht durchsetzen können. Beispielsweise führt die Geschlechtsdifferenzierung der neuen Europabevölkerung in drei Bevölkerungsstrukturen (weiblich, männlich, insgesamt) dazu, dass die altersstandardisierten Mortalitätsraten von Männern und Frauen nicht mehr vergleichbar sind, weil sie mit unterschiedlichen Standardbevölkerungen gewichtet wurden. In der Gesundheitsberichterstattung der Länder wird daher wieder auf die „alte“ Europabevölkerung zurückgegriffen und somit sichergestellt, dass eine Vergleichbarkeit der standardisierten Raten auch zwischen den Geschlechtern gegeben ist.

Alte und neue Europa-Standard-Bevölkerung

Alter von ... bis ... Jahren	"Alte" Europa- Standard- Bevölkerung	"Neue" Europa-Standard-Bevölkerung		
		männlich	weiblich	insgesamt
0 - 1	1 600	1 345	1 218	1 305
1 - 4	6 400	5 303	4 800	5 021
5 - 9	7 000	6 800	6 160	6 472
10 - 14	7 000	7 108	6 452	6 772
15 - 19	7 000	7 570	6 863	7 208
20 - 24	7 000	8 163	7 438	7 792
25 - 29	7 000	8 206	7 552	7 871
30 - 34	7 000	7 811	7 258	7 528
35 - 39	7 000	7 448	6 986	7 212
40 - 44	7 000	7 068	6 661	6 860
45 - 49	7 000	5 997	5 739	5 865
50 - 54	7 000	5 937	5 817	5 876
55 - 59	6 000	5 521	5 585	5 553
60 - 64	5 000	5 015	5 463	5 245
65 - 69	4 000	4 139	5 196	4 680
70 - 74	3 000	2 449	3 392	2 932
75 - 79	2 000	2 228	3 536	2 897
80 - 84	1 000	1 094	2 076	1 606
85 u. mehr	1 000	798	1 808	1 305
Insgesamt	100 000	100 000	100 000	100 000

Quelle:
Cancer incidence in five
continents, Lyon, IARC, 1976, Vol.3

nach:
United Nations World Population Prospects 1990

Indirekte Altersstandardisierung

Die Methode der indirekten Altersstandardisierung geht bei der Gewichtung genau umgekehrt vor wie die direkte Standardisierung. Es werden nun die altersspezifischen Sterberaten der Standardbevölkerung mit der Altersverteilung der untersuchten Population gewichtet. Es resultiert im ersten Rechenschritt die Anzahl der Sterbefälle, die man in der untersuchten Bevölkerung erwarten würde, wenn das Sterbegeschehen gleich dem in der Standardbevölkerung wäre:

$$\text{Erwartete Sterbefälle} = \sum \frac{n_i \cdot MR_i}{100\,000}$$

Dabei steht n_i für die Zahl der Personen in der Altersgruppe i der beobachteten Bevölkerung und MR_i für die Mortalitätsrate je 100.000 der Altersgruppe i in der Standardbevölkerung.

Die erwarteten Sterbefälle, die sich aus der Summe über alle Altersgruppen ergeben, werden dann im zweiten Schritt in Beziehung gesetzt zu den tatsächlich in der untersuchten Bevölkerung aufgetretenen Todesfällen: Die Division der beobachteten durch die erwarteten Fälle führt zur so genannten Standardized Mortality Ratio (SMR).

$$SMR = \frac{\text{Anzahl beobachteter Fälle}}{\text{Anzahl erwarteter Fälle}} = \frac{d}{\sum \frac{n_i \cdot MR_i}{100\,000}}$$

wobei d die Summe aller Sterbefälle in der untersuchten Population angibt. In der deutschsprachigen Literatur sind die Begriffe standardisierte Mortalitätsratio bzw. Mortalitätsverhältniszahl und standardisierter Mortalitätsindex im Gebrauch. Die Übersetzung „standardisierte Mortalitätsrate“ ist nicht zutreffend, da es sich um einen Ratenquotienten handelt.

In einem dritten Schritt kann man nun die SMR mit der rohen Sterberate der Standardpopulation multiplizieren und erhält die

$$\text{Indirekt standardisierte Mortalitätsrate } MR_{is} = MR \cdot SMR$$

Allerdings interessiert bei der indirekten Standardisierung in erster Linie der SMR-Quotient. Üblicherweise wählt man bei dieser Methode die Summe aller untersuchten Regionen als Standardpopulation; für die bayerischen Landkreise und kreisfreien Städte wäre das sinnvollerweise die Bevölkerung des Landes Bayern. Die SMR des Landes insgesamt ist dann definitionsgemäß exakt 1 (beobachtete = erwartete Fälle). Die Ergebnisse für die einzelnen Landkreise bzw. kreisfreien Städte lassen sich als prozentuale Abweichung von diesem Landesdurchschnitt interpretieren und können auf statistische Signifikanz getestet werden (siehe dazu den nächsten Abschnitt). Anders als die direkte Altersstandardisierung erlaubt die indirekte Standardisierung immer nur den Zweier-Vergleich zwischen der untersuchten Population und der Standardpopulation: die SMR hängt vom Altersaufbau der untersuchten Population ab und ist um den Altersaufbau der Standardpopulation bereinigt. Daher können anhand der SMR nicht Mortalitätsunterschiede *mehrerer* Landkreise und kreisfreier Städte *untereinander* verglichen werden.

Vergleich mit dem Landesdurchschnitt

Man benötigt für die indirekte Altersstandardisierung nur den Altersaufbau und die Gesamtzahl der Todesfälle in der beobachteten Bevölkerung, während in die direkte Standardisierung die altersspezifischen Sterberaten dieser Population mit eingehen. Bei der indirekten Altersstandardisierung verzichtet man also darauf, die Information über die altersspezifische Sterblichkeit in der untersuchten Population zu nutzen. Das ist vor allem dann sinnvoll, wenn es um die Untersuchung von kleineren Grundgesamtheiten mit geringen Fallzahlen geht. Wenn beispielsweise an einer seltenen Todesursache in einer Untersuchungsregion nur 10 Personen pro Jahr versterben, verteilen sich die Sterbefälle recht zufällig auf die einzelnen Altersgruppen, so dass keine verlässlichen altersspezifischen Mortalitätsraten erwartet werden können. Führt man in dieser Situation eine direkte Altersstandardisierung durch, so erhalten diese Zufallsschwankungen einen unzulässig großen Einfluss auf das Gesamtergebnis, während die indirekte Methode hier durch die Verwendung der zuverlässigeren Gesamtfallzahl den Fehler gering hält und stabilere Schätzungen erbringt.

Kleine Fallzahlen: indirekte Standardisierung



SMR-Quotienten, die anhand der Mortalitätsrate Bayerns berechnet werden, sind nützlich, um zu beurteilen, ob die Sterblichkeit in einem Landkreis bzw. einer kreisfreien Stadt vom Landesdurchschnitt abweicht. Die so berechneten SMR-Quotienten lassen aber weder einen Vergleich mit anderen Bundesländern zu noch dürfen mehrere SMR-Quotienten untereinander verglichen werden.

Um die Ergebnisse der Altersstandardisierung statistisch stabiler zu machen, empfiehlt es sich, vor allem bei der Betrachtung der Sterblichkeit nach einzelnen Todesursachen, die Daten mehrerer Jahre zusammenzufassen. In Anlehnung an das Verfahren in Nordrhein-Westfalen wird die Landes-Gesundheitsberichterstattung in Bayern für Vergleiche zwischen Landkreisen und kreisfreien Städten die Daten dreier Jahre zusammenfassen.

Jahresmittelwerte

3.2. Signifikanzprüfung

Wenn man Kenngrößen der Gesundheitsstatistik wie z. B. Sterblichkeitsraten einem regionalen Vergleich unterzieht, ist für die sinnvolle Interpretation von Differenzen zwischen einzelnen Kreisen oder von Abweichungen gegenüber dem Landesdurchschnitt eine statistische Signifikanzprüfung der Differenzen unerlässlich. Der Signifikanztest überprüft, ob Datendifferenzen im Bereich zufälliger Schwankungen liegen oder aber mit großer Wahrscheinlichkeit einen tatsächlichen Unterschied anzeigen. Dabei wird eine sog. „Irrtumswahrscheinlichkeit“ festgesetzt, die angibt, mit welcher Wahrscheinlichkeit die getesteten Werte, wären sie denn zufällig zustande gekommen, dennoch den Signifikanztest bestehen würden. Als Irrtumswahrscheinlichkeit reicht für die Gesundheitsberichterstattung in der Regel ein Wert von 5 % bzw. von 1 % aus. Die zu 100 % komplementären Werte 95 % bzw. 99 % werden „Vertrauenswahrscheinlichkeit“ genannt.

Es gibt verschiedene statistische Methoden zur Signifikanzprüfung, wobei sich der Signifikanztest auf der Basis von Konfidenzintervallen (Vertrauensbereichen) für die meisten Zwecke der Gesundheitsberichterstattung anbietet. Im Folgenden soll jeweils eine Formel zur Berechnung von Konfidenzintervallen für direkt altersstandardisierte Raten und für den SMR-Quotienten vorgestellt werden.

Konfidenz- oder Vertrauensintervall

Konfidenzintervalle bei direkter Altersstandardisierung

Vorausgesetzt, die Zahl der Todesfälle ist groß genug (25 kann hier als untere Grenze gelten) ergibt sich das 95 % - Konfidenzintervall der altersstandardisierten Mortalitätsrate MR_{st} als

$$CI_{95\%} = MR_{st} \pm 1,96 \cdot \sqrt{\sum \frac{mr_i \cdot (100\,000 - mr_i)}{n_i} \cdot \left(\frac{N_i}{\sum N_i} \right)^2}$$

wobei das "+" Zeichen für die obere und das "-" Zeichen für die untere Grenze des Konfidenzintervalls einzusetzen sind. Der Wert 1,96 ergibt sich aus den Perzentilen der Normalverteilung. Für ein Konfidenzniveau von 99 % wäre hier statt dessen der Wert 2,576 einzusetzen.

Konfidenzintervalle bei indirekter Altersstandardisierung

Die Berechnung des 95 %-Konfidenzintervalls für die „Standardized Mortality Ratio“ (SMR) ergibt sich wie folgt. Die SMR wird, wie oben erläutert, als Quotient aus beobachteten und erwarteten Fällen gebildet:

$$SMR = \frac{\text{Anzahl beobachteter Fälle}}{\text{Anzahl erwarteter Fälle}} = \frac{O}{E}$$

Unter der Annahme, dass die beobachteten Todesfälle einer Poissonverteilung genügen, lässt sich für den SMR-Quotienten der untere und der obere Grenzwert des 95 %-Konfidenzintervalls vereinfacht folgendermaßen berechnen:

$$CI_{95\%u} = \frac{\left(\frac{1,96}{2} - \sqrt{O}\right)^2}{E} \quad \text{und} \quad CI_{95\%o} = \frac{\left(\frac{1,96}{2} + \sqrt{O+1}\right)^2}{E}$$

Allerdings sollte diese Näherungsformel erst ab einer Fallzahl über 100 verwendet werden. Ein 95 %-Konfidenzintervall zur SMR für Fallzahlen ab 30 liefert die folgende Formel:

$$CI_{95\%} = \frac{O + 0,5 \cdot (1,96)^2 \pm 0,5 \pm 1,96 \cdot \sqrt{O + 0,25 \cdot (1,96)^2 \pm 0,5}}{E}$$

Mit den ermittelten Konfidenzintervallen lässt sich nun die Differenz zweier Raten ganz einfach auf Signifikanz testen: Signifikante Abweichungen liegen immer dann vor, wenn sich die Vertrauensbereiche der zu vergleichenden Raten nicht überschneiden.

3.3. Weitere Verfahren zur Darstellung der Sterblichkeit

In der Gesundheitsberichterstattung hat in den letzten Jahren eine Reihe weiterer Verfahren zur Bearbeitung von Sterblichkeitsziffern Verbreitung gefunden. So wird beispielsweise neben der Gesamtsterblichkeit auch die „**Vermeidbare Sterblichkeit**“ berechnet. Dabei wird für ausgewählte Todesursachen eine Altersgrenze angenommen, unterhalb derer die Todesfälle als grundsätzlich vermeidbar betrachtet werden. Meist liegt diese Grenze bei 65 Jahren. Die vermeidbare Sterblichkeit wird für diese Todesursachen in den definierten Altersgrenzen in gleicher Weise berechnet und standardisiert wie die Gesamtsterblichkeit.

**Konfidenzintervall
für große
Fallzahlen**

**Konfidenzintervall
für kleine
Fallzahlen**

**Vermeidbare
Sterblichkeit**

Einige zusätzliche Rechenschritte sind notwendig für die Berechnung des Gesundheitsindikators „**Verlorene Lebensjahre**“ (PYLL - Potential Years of Life Lost). Auch er ist ein Maß für die vorzeitige Sterblichkeit und wird üblicherweise für die Altersgruppe der 1- bis 64-Jährigen berechnet. Ausgehend von der vorgegebenen Altersobergrenze werden für alle vor Erreichung der Altersgrenze Verstorbenen die nicht gelebten Jahre, das verlorene Potenzial, aufsummiert sowie als Rate je 100.000 Einwohner gebildet. Die absolute Zahl der „Verlorenen Lebensjahre“ (VLJ) lässt sich jedoch relativ schwer interpretieren. Von größerem Interesse ist daher der prozentuale Anteil verschiedener Todesursachen an der Gesamtzahl der VLJ, der es erlaubt, das präventive Potenzial der wichtigsten Ursachen für vorzeitigen Tod in unserer Gesellschaft zu bewerten.

Die Berechnung der VLJ für eine bestimmte Todesursache bzw. Ursachengruppe beruht auf der Summierung der betreffenden Sterbefälle in den einzelnen Altersgruppen, multipliziert mit den in der jeweiligen Altersgruppe verbleibenden Lebensjahre bis zur vorgegebenen oberen Altersgrenze. Der Mittelwert verlorener Lebensjahre je Todesfall in einer Altersklasse VLJ_{m-i} ergibt sich bei einem Grenzwert von 65 Jahren (gestorben vor Vollendung des 65. Lebensjahres) wie folgt:

$$VLJ_{m-i} = 65 - \left(\frac{UG_i + OG_i}{2} \right) - 0,5$$

wobei UG_i und OG_i die unteren und oberen Grenzwerte der jeweiligen Altersgruppe sind. Für die Altersgruppe der 40- bis 44-Jährigen ergäbe sich somit ein mittlerer Verlust von 22,5 Lebensjahren durch Tod vor Vollendung des 65. Lebensjahres. Die Formel zur Berechnung der VLJ (zwischen 1 und 64 Jahren) lautet dann:

$$VLJ_{(1-64)} = \sum_{i=1}^{64} d_i \cdot \left(65 - \frac{UG_i + OG_i}{2} - 0,5 \right)$$

Dabei steht hier d_i für die Anzahl der Sterbefälle in der Altersgruppe i .

Die Gruppe der Säuglinge wird bei dieser Berechnung ausgeklammert, da die Säuglingssterblichkeit zum größten Teil speziellen Todesursachen zuzuschreiben ist, die in späteren Lebensabschnitten nicht mehr auftreten. Da die Säuglinge in der unter 55jährigen Bevölkerung die höchste Sterblichkeitsrate aufweisen und jeder Sterbefall 64 VLJ entspricht, würden die speziellen Diagnosen der Neo- und Postneonatalperiode die Statistik der VLJ sehr stark beeinflussen.

Die VLJ werden in der Regel als Rate je 100.000 Einwohner angegeben und lassen sich mit der weiter oben vorgestellten Methode altersstandardisieren:

$$VLJ_{(1-64)st} = \frac{\sum_{i=1}^{64} d_i \cdot \left(65 - \frac{UG_i + OG_i}{2} - 0,5 \right) \cdot \frac{N_i}{n_i}}{\sum_{i=1}^{64} N_i} \cdot 100\,000$$

Es existieren weitere, komplexe Verfahren der Berechnung von VLJ auf der Basis von Sterbetafeln, allerdings weichen die hiermit gewonnenen Ergebnisse nur geringfügig von denen der einfachen Methode ab.

Eine weitere wichtige Maßzahl in der Gesundheitsberichterstattung ist die **Lebenserwartung**. Die Lebenserwartung ist eine Maßzahl, die modellhaft aus den alters- und geschlechtsspezifischen Sterblichkeitsverhältnissen einer Bevölkerung eine erwartete Lebensdauer ableitet. Sie gibt an, wie viele Jahre ein Mensch in einem definierten Alter noch leben würde, wenn die gegenwärtigen altersspezifischen Mortalitätsraten für den Rest seines Lebens unverändert blieben. Die fernere Lebenserwartung der Nulljährigen wird hierbei auch als mittlere bzw. durchschnittliche Lebenserwartung bezeichnet. Die mittlere Lebenserwartung ist als eine Momentaufnahme anzusehen, die in komprimierter Form die aktuellen Sterblichkeitsverhältnisse und somit indirekt die gesundheitliche Lage der untersuchten Bevölkerung wiedergibt. Für ihre Berechnung wird keine Standardbevölkerung benötigt und sie ist für den Laien deutlich anschaulicher als andere Mortalitätskennziffern. Die Berechnung der Lebenserwartung soll hier allerdings nicht dargestellt werden – sie stellt etwas anspruchsvollere Voraussetzungen an die verfügbaren Daten.

Lebenserwartung



Formeln ohne Ende? Hilfestellungen für Landkreise und kreisfreie Städte in Bayern

Die Arbeitsgemeinschaft der Obersten Landesgesundheitsbehörden hat 2003 eine CD-Rom mit dem neuen Indikatorenset für die Gesundheitsberichterstattung der Länder herausgegeben. Diese CD-Rom enthält auch Excel-Arbeitsblätter mit einprogrammierten Formeln, mit deren Hilfe Landkreise und kreisfreie Städte recht einfach altersstandardisierte Sterblichkeiten, PYLL und Lebenserwartung berechnen können. Das Bayerische Landesamt für Gesundheit und Lebensmittelsicherheit stellt diese Arbeitsblätter auf Anfrage zur Verfügung und führt auf Anfrage auch Schulungen zur Berechnung einzelner Indikatoren durch.

Ansprechpartner:

Joseph Kuhn
Bayerisches Landesamt für
Gesundheit und Lebensmittelsicherheit
Veterinärstr. 2, 85764 Oberschleißheim
Tel.: 089/31560-302
Mail: joseph.kuhn@lgl.bayern.de

Hilfestellung

4. Epidemiologische Studien

4.1. Studientypen: ein kurzer Überblick

Querschnittsstudien

Der Begriff „Querschnittsstudie“ wird oft synonym mit Prävalenzstudie oder Cross-sectional-study verwendet. In Querschnittsstudien wird zu einem einzigen Zeitpunkt eine Population (oder eine Stichprobe daraus) untersucht. Es werden verschiedene Merkmale erfasst und korrelativ miteinander in Beziehung gesetzt. Für den kausalen Nachweis eines Zusammenhangs von Expositionen und Krankheiten sind Querschnittsstudien in der Regel wenig geeignet. Sie eignen sich vor allem für die Erhebung von Prävalenzen und für die Bildung von Hypothesen über mögliche Krankheitsursachen. Querschnittsstudien sind relativ einfach und ökonomisch durchzuführen.

**Querschnitts-
studien**

Eine typische Querschnittsstudie im Gesundheitsamt ist die Schulinganguntersuchung. Auch die vom Landesamt für Statistik und Datenverarbeitung veröffentlichten Jahresscheiben der Todesursachenstatistik haben den Charakter einer Querschnittserhebung. In der Gesundheitsberichterstattung werden häufig Daten aus solchen Querschnitt-Designs verwendet, da in der Gesundheitsberichterstattung das Monitoring-Ziel, also das Beobachten der gesundheitlichen Verhältnisse, gegenüber dem Forschungs-Ziel, also dem Auffinden neuer Zusammenhänge, im Vordergrund steht.

Fall-Kontroll-Studien

Synonyme dafür sind case-control-study oder case-comparison-study. Bei Fall-Kontroll-Studien wird eine Gruppe von Erkrankten und Nicht-Erkrankten im Hinblick auf eine zurückliegende (ggf. hypothetische) Exposition untersucht. Das Verhältnis von Gesunden und Kranken ist hier also durch die Auswahl der Studienpopulation festgelegt. Man kann daher keine Inzidenzen und Relativen Risiken berechnen, sondern greift auf die Berechnung von Odds Ratios zurück. Fall-Kontroll-Studien eignen sich besonders zur Untersuchung der Ursachen seltener Erkrankungen. Wichtig bei Fall-Kontroll-Studien ist, dass die beiden Gruppen, die sog. „Fälle“ und „Kontrollen“, vergleichbar und möglichst repräsentativ für die Population sind, über die man Aussagen treffen möchte.

**Fall-Kontroll-
Studien**

Kohortenstudien

Synonyme dafür sind Longitudinalstudien, Follow-Up-Studien oder prospektive Studien. In gewisser Weise arbeiten Kohortenstudien umgekehrt wie die Fall-Kontrollstudien. Es wird eine gesunde Population, bestehend aus einer Gruppe von Exponierten und einer Gruppe von Nichtexponierten, in ihrer weiteren Entwicklung beobachtet, das Auftreten von Erkrankungen in den beiden Gruppen registriert und analysiert. Kohortenstudien sind gut geeignet, um die (ggf. auch unbekannt) gesundheitlichen Folgen von (bekannt) Expositionen zu analysieren. Sie lassen die Berechnung von Inzidenzen und somit Relativen Risiken zu.

**Kohorten-
studien**

Ein typisches Problem bei Kohortenstudien besteht darin, dass die Kohorten oft nicht über den gesamten Untersuchungszeitraum konstant bleiben, z.B. weil Personen sterben oder wegziehen.

Eine Sonderform der Kohortenstudie ist die retrospektive Kohortenstudie, bei der man Daten über eine in der Vergangenheit gesunde Population hat.

Experimentelle Studien

Experimentelle Studien, kontrollierte Studien oder Interventionsstudien zeichnen sich dadurch aus, dass – im klassischen Design – die Einflussfaktoren experimentell variiert werden. Es gibt daneben auch ein sog. „quasi-experimentelles Design“, bei dem die Variation der Einflussfaktoren durch externe Gegebenheiten erfolgt und nur beobachtet wird. Inzidenzen und Relative Risiken können berechnet werden. Experimentelle Studien sind sehr gut für den Nachweis kausaler Faktoren geeignet. Allerdings ist es aus technischen oder ethischen Gründen nicht immer möglich, die fraglichen Expositionen einem experimentellen Design zu unterwerfen.

Kontrollierte Studien

Von einer „Einfach-Blind-Studie“ spricht man, wenn nur die untersuchte Person nicht weiß, welcher Exposition sie ausgesetzt ist (z.B. echtes Medikament oder Placebo), bei einer „Doppel-Blind-Studie“ weiß das auch die Studienleitung nicht. Damit sollen z.B. sozialpsychologische Einflüsse auf das Studienergebnis (Erwartung und Suggestion von Effekten usw.) möglichst ausgeschlossen werden. Erfolgt die Verteilung der Personen danach, ob sie das Medikament oder das Placebo bekommen, nach einem Zufallsverfahren, spricht man von einer „randomisierten Studie“ (RCT – Randomized Clinical Trial). Wenn Medikament und Placebo auf unterschiedliche Personen verteilt werden, handelt es sich um ein „Parallel-Design“, wenn die gleichen Personen einmal das Medikament und dann das Placebo bekommen, um ein „Crossover-Design“ der Studie.

4.2. Die Qualität von Studien

Leitlinien guter epidemiologischer Praxis

Bei der Beurteilung der Qualität von Studien spielen vielfältige ethische, inhaltliche, statistische, dokumentationstechnische und ökonomische Aspekte eine Rolle. Die wichtigsten sind in den „Leitlinien und Empfehlungen zur Sicherung Guter Epidemiologischer Praxis (GEP)“ zusammengefasst (DAE 2000, S. 295 ff., im Internet unter http://www.gmds.de/download/empf_ gep_langfassung.pdf):

Leitlinien guter epidemiologischer Praxis

- Epidemiologische Studien müssen Menschenwürde und Menschenrechte respektieren; die Stellungnahme einer Ethikkommission soll vorab eingeholt werden.
- Der Studienplanung sollen explizite Forschungsfragen mit einem detaillierten Forschungsplan zugrunde liegen. Darin sind z.B. die Ausgangshypothesen, das Studiendesign, die Auswahl der Studienpopulation usw. festzuhalten.

- Bei Studien mit biologischen Proben ist vorab festzulegen, was mit den Proben nach der Studie geschieht; das Einverständnis der Probanden ist einzuholen.
- In der Studiendurchführung sollen Verfahren der Qualitätssicherung angewandt werden (Dokumentation von Veränderungen des Studiendesigns etc.).
- Die Erhebung und Verarbeitung aller Daten soll detailliert dokumentiert werden. Rohdatensätze und bearbeitete Datensätze sind zu trennen.
- Die Auswertung der Daten soll in nachvollziehbarer Weise erfolgen und es sollen Gegenprüfungen ermöglicht werden.
- Der Datenschutz ist zu gewährleisten.
- Die rechtlichen Rahmenbedingungen, z.B. zwischen Auftraggeber und Auftragnehmer, sollen klar geregelt werden. Das schließt Fragen der Datenverwendung und der Publikationsrechte ein.
- Studieninstrumente und -ergebnisse sollen nach Möglichkeit öffentlich zugänglich sein.

Zudem verpflichtet das ärztliche Berufsrecht grundsätzlich dazu, vor Durchführung einer epidemiologischen Studie das Vorhaben der zuständigen Ethikkommission (z.B. Ethikkommission der Bayerischen Landesärztekammer, www.blaek.de) zur Beratung vorzulegen.

Ethikkommission



Unter der Lupe: Methodische Qualitätsmängel

Ein spezieller Aspekt der Qualität von Studien betrifft methodische Fehler bei der Generierung, Auswertung und Interpretation der Daten. Einige typische Fehler sollen hier kurz vorgestellt werden:

- **Selektions-Bias:** Unter einem Bias wird die systematische Verzerrung in den Daten verstanden. Der Selektions-Bias bezeichnet einen Stichprobenfehler, durch den ungewollt eine Krankheit über- oder unterschätzt wird. Ein bekanntes Beispiel ist der „healthy worker effect“, den man als Möglichkeit bedenken muss, wenn man die Häufigkeit einer Krankheit in einem Betrieb untersuchen will: eventuell sind nämlich die ernsthaft Erkrankten schon aus dem Arbeitsprozess ausgeschieden.
- **Response-Bias:** Auch dies ist ein Stichprobenfehler, der Verzerrungen bei der Beteiligung an Befragungen bezeichnet. Beispielsweise kann es sein, dass bei einer Befragung von Jugendlichen in einer Schule zum Thema „Rauchen“ eher die Nichtraucher antworten.
- **Interviewer-Bias:** Eine Untersuchung, vor allem eine Befragung, ist immer ein sozialer Prozess. Seitens des Interviewers können Antworten ungewollt suggestiv beeinflusst werden. Umgekehrt neigen Befragte möglicherweise dazu, in eine sozial erwünschte Richtung zu antworten.

**Bias,
systematische
Verzerrung**

- **Fallzahlenproblem:** Um Krankheitshäufigkeiten und vor allem Unterschiede zwischen einzelnen Teilpopulationen verlässlich zu erfassen, müssen die Stichproben groß genug sein. Während in wissenschaftlichen Studien in aller Regel die notwendige Stichprobengröße vorab berechnet wird, ist das bei vielen Befragungsaktionen von Praktikern nicht der Fall. Die Verlässlichkeit der Aussagen muss dann ggf. durch geeignete statistische Verfahren, z.B. die Berechnung von Konfidenzintervallen, nachgeprüft werden, bevor man die Ergebnisse übernimmt.
- **Verallgemeinerungsfehler:** Häufig wird von Ergebnissen, die an speziellen Untersuchungsgruppen gewonnen wurden, unzulässigerweise auf die Gesamtbevölkerung geschlossen. Beispielsweise ließ die amerikanische Food and Drug Administration bis Ende der 80er Jahre Studien mit neuen Medikamenten nur an Männern durchführen. Die Ergebnisse wurden einfach auf die Frauen übertragen. Das umgekehrte Problem, nämlich die unreflektierte Übertragung von statistischen Zusammenhängen, die an einer Gruppe gewonnen wurden, auf Einzelpersonen oder Teilgruppen, nennt man einen „Ökologischen Fehlschluss“.
- **Confounding:** Merkmale, die statistisch korrelieren, täuschen leicht kausale Zusammenhänge vor. So gibt es z.B. statistische Zusammenhänge zwischen Schlaganfall und Demenz sowie zwischen Herzinfarkt und Demenz. Während Schlaganfälle tatsächlich zur (vaskulären) Demenz führen können, haben Herzinfarkte darauf in der Regel keinen Einfluss, aber beide Erkrankungen treten mit zunehmendem Alter häufiger auf. Einen solchen Faktor, der einen statistischen Zusammenhang von eigentlich unabhängigen Merkmalen erzeugt, nennt man Confounder.

Fallzahlen

Verallgemeinerung

Confounding

4.3. Evidenzklassen

Medizinische Handlungsempfehlungen werden seit einiger Zeit im Rahmen des Ansatzes der „Evidence based Medicine“ danach bewertet, welche Studientypen in welcher Qualität ihnen zugrunde liegen. Dazu werden „Evidenzklassen“ gebildet. Es gibt verschiedene Formen der Einteilung der Evidenzstärke in Evidenzklassen. Die folgende Einteilung nach Evidenzklassen stammt vom Scottish Intercollegiate Guideline Network (SIGN) und wird auch von der Zentralstelle der Deutschen Ärzteschaft für Qualitätssicherung in der Medizin (eine gemeinsame Einrichtung der Bundesärztekammer und der Kassenärztlichen Bundesvereinigung) angewandt.

Evidenzklassen	
1 ⁺⁺	Hochwertige Metaanalysen, systematische Reviews zu RCTs oder RCTs mit einem sehr geringen Risiko von Verzerrungen (Bias)
1 ⁺	Gut durchgeführte Metaanalysen, systematische Reviews zu RCTs oder RCTs mit einem geringen Risiko von Verzerrungen (Bias)
1 ⁻	Metaanalysen, systematische Reviews zu RCTs oder RCTs mit einem hohen Risiko von Verzerrungen (Bias)
2 ⁺⁺	Hochwertige systematische Reviews zu Fall-Kontroll- oder Kohortenstudien Hochwertige Fall-Kontroll- oder Kohortenstudien mit einem sehr geringen Risiko von Verzerrungen (Confounder, Bias) oder Zufall und einer hohen Wahrscheinlichkeit, dass Kausalität besteht
2 ⁺	Gut durchgeführte Fall-Kontroll- oder Kohortenstudien mit einem geringen Risiko von Verzerrungen (Confounder, Bias) oder Zufall und einer mittleren Wahrscheinlichkeit, dass Kausalität besteht
2 ⁻	Fall-Kontroll- oder Kohortenstudien mit einem hohen Risiko von Verzerrungen (Confounding, Bias) oder Zufall und einem signifikanten Risiko, dass keine Kausalität besteht
3	Nicht-analytische Studien, z.B. Fallberichte, Fallserien
4	Expertenmeinung
Quelle: http://www.leitlinien.de/versorgungsleitlinien/VLEvidenz/view , Download vom 30.1.2004.	

Solche Evidenzklasseneinteilungen sind hilfreich, wenn damit z.B. Behandlungsleitlinien oder Präventionsmaßnahmen bewertet werden. Es geht dabei im Kern immer um die Beurteilung der Kausalität von Zusammenhängen bzw. der Wirksamkeit von Interventionen.

4.4. Objektivität, Validität, Reliabilität, Sensitivität und Spezifität

Bei Angaben zur Qualität epidemiologischer Studien ist häufig die Rede von der Objektivität, Validität, Reliabilität, Sensitivität und Spezifität der eingesetzten Testverfahren bzw. Untersuchungsinstrumente.

Die **Objektivität** eines Verfahrens ist gegeben, wenn seine Ergebnisse unabhängig vom Untersucher sind. Eine Einschränkung der Objektivität ist häufig bei nicht vollständig standardisierten Verfahren

gegeben: das Ergebnis der Untersuchung hängt dann zu einem gewissen Grad vom „Ermessen“ des Untersuchers ab.

Testverfahren

Die **Validität** (Gültigkeit) meint die inhaltliche Genauigkeit, mit der ein Instrument misst, was es messen soll. Das ist vor allem bei Instrumenten, mit denen psychische Störungen oder Verhaltensstörungen erfasst werden sollen, häufig ein Problem. Die messtheoretischen Voraussetzungen für deren valide quantitative Erfassung sind nicht selten fragwürdig. Während man einigermaßen genau definieren kann, was unter einer Masernerkrankung oder einem Schlaganfall zu verstehen ist und solche Festlegungen auch relativ unbelastet von theoretischen Vorannahmen sind, ist das z.B. beim Aufmerksamkeits-Defizit-Syndrom (ADHS) ganz anders. Das Konstrukt selbst ist theoretisch umstritten, die Diagnosestellung schwierig und die eingesetzten Testverfahren möglicherweise unterschiedlich valide, d.h. sie messen unter Umständen ganz verschiedene Dinge. Testverfahren für solche Störungsbilder müssen daher validiert werden, z.B. durch Korrelation mit einem anderen, anerkannten, Testverfahren oder mit einem unstrittigen Außenkriterium.

Die **Reliabilität** (Zuverlässigkeit) hebt auf die Wiederholungsgenauigkeit eines Verfahrens ab. Ein Instrument ist dann reliabel, wenn die Ergebnisse einer ersten Durchführung gut mit denen einer zweiten Durchführung korrelieren.

Für die Epidemiologie von besonderer Bedeutung sind die **Sensitivität** und **Spezifität** eines Verfahrens. Dabei geht es einmal darum, in welchem Umfang ein Test die Erkrankungsfälle erkennt (Sensitivität), und einmal darum, in welchem Umfang er die gesunden Personen erkennt (Spezifität).

Zur Veranschaulichung dieser Maße kann man wiederum eine Vierfeldertafel benutzen. Die Werte in den Zellen müssen durch aufwändige Validierungsstudien für den jeweiligen Test gewonnen werden, also aus der genauen Diagnostik der test-positiven und der test-negativen Fälle.

	Erkrankt	Nicht erkrankt	Gesamt
Im Test positiv	a	b (falsch positiv)	a + b
Im Test negativ	c (falsch negativ)	d	c + d
Gesamt	a + c	b + d	

$$\text{Sensitivität} = \frac{a}{a + c}$$

$$\text{Spezifität} = \frac{d}{b + d}$$

Sind Sensitivität und Spezifität eines Tests aus Validierungsstudien bekannt, so kann man mit der gleichen Vierfeldertafel, wenn man die tatsächliche Prävalenz der Erkrankung in der Bevölkerung (bzw. in der Untersuchungspopulation) kennt, den sog. „**positiven prädiktiven Wert**“ des Tests berechnen. Dieser Wert gibt an, mit welcher Wahrscheinlichkeit ein positiv getesteter Mensch tatsächlich krank ist:

$$\text{Positiver prädiktiver Wert} = \frac{a}{a + b}$$

Diese Formel gilt jedoch nur, wenn die Randzellen die tatsächlichen Verhältnisse in der Bevölkerung bzw. in der Untersuchungspopulation wiedergeben.

Für die Gesundheitsberichterstattung ist die Qualität von Testverfahren, wie sie z.B. in Screenings wie der Schuleingangsuntersuchung angewandt werden, vor allem im Hinblick auf die Verlässlichkeit der gewonnenen Daten relevant. Abschließend sei noch auf einen anderen Aspekt hingewiesen: Im öffentlichen Gesundheitsdienst werden Testverfahren primär nicht zur Gewinnung von Daten für die Gesundheitsberichterstattung eingesetzt, sondern um Einzelfallentscheidungen zu begründen, die den weiteren Lebensweg der getesteten Menschen mitunter erheblich beeinflussen. Mit anderen Worten: Die Aussagekraft der Testverfahren beurteilen zu können, z.B. den positiven prädiktiven Wert eines Tests, ist für eine qualifizierte und verantwortliche Beratung der getesteten Personen von ausschlaggebender Bedeutung. Auch bei einem Testverfahren mit zufriedenstellender Sensitivität und Spezifität treten nämlich viele falsch positive Fälle auf, wenn die Prävalenz der Erkrankung in der Gruppe, in der man testet, niedrig ist. Das ist häufig bei bevölkerungsbezogenen Screenings der Fall. Personen mit einem positiven Testergebnis z.B. auf Brustkrebs oder AIDS müssen dann qualifiziert über die Bedeutung dieses Ergebnisses und die Notwendigkeit einer genauen Abklärung des Screenings-Befundes beraten werden.

Vorhersagekraft

**Testqualität
und Beratung**



Ein Beispiel: In der Zeitschrift „Kinderärztliche Praxis“ 5/2003 wird ein in vielen Anwendungen bewährtes Legastheniescreening für seine Tauglichkeit in bevölkerungsbezogenen Studien bewertet. Die Screeningautoren geben die Sensitivität des Verfahrens mit 64 % und die Spezifität mit 80 % an. Der Herausgeber der Zeitschrift

geht nun von einer Prävalenz der Legasthenie von 5 % aus und errechnet damit für den Test einen positiven prädiktiven Wert von 14 %. Das heißt: nur 14 % der im Test positiven Kinder haben wirklich eine Legasthenie. Die Berechnung des positiven prädiktiven Werts ergibt sich aus der Vierfeldertafel, wenn man die Zellen mit den angegebenen Werten füllt:

	Erkrankt	Nicht erkrankt	Gesamt
Im Test positiv	32 (64 % von 50)	190	222 (32 + 190)
Im Test negativ	18	760 (80 % von 950)	778 (18 + 760)
Gesamt	50 (Prävalenz)	950	1000 (fiktive Grund- gesamtheit)

**Tests im
Screening**

Der Herausgeber der Zeitschrift zieht daraus den Schluss, dass ein solches Verfahren für ein *bevölkerungsbezogenes* Screening nicht sinnvoll ist. Auch dazu eine Berechnung: Das Verfahren testet 222 von 1.000 Kindern positiv. Bei den Schuleingangsuntersuchungen in Bayern mit ca. 120.000 Kindern wären das fast 27.000 Kinder, für die eine Abklärung der Verdachtsdiagnose aus dem Screening notwendig wäre. Für ca. 23.000 Kinder würde die Abklärung einen negativen Befund erbringen – eine hohe Sensitivität und Spezifität der Nachuntersuchung vorausgesetzt.

Ganz anders verhält es sich, wenn das Screening auf eine selektierte Gruppe angewandt wird, z.B. auf schulauffällige Kinder. In dieser Gruppe ist die Prävalenz höher als in der Gesamtbevölkerung, dadurch steigt auch der positive prädiktive Wert des Screenings für diese Kinder.

5. Literatur



Literaturtipps zur Darstellung epidemiologischer Ergebnisse in der Gesundheitsberichterstattung:

1. Gigerenzer, G.: Das Einmaleins der Skepsis. Über den richtigen Umgang mit Zahlen und Risiken. Berlin 2002. Dieses Buch zeigt, wie sehr die Verständlichkeit epidemiologischer Sachverhalte von ihrer Darstellungsform abhängt – für die Gesundheitsberichterstattung eine empfehlenswerte Lektüre.

2. Beck-Bornholdt, H.-P., Dubben, H.-H.: Der Hund, der Eier legt. Reinbek bei Hamburg, 2001. Eine amüsante Einführung in die Fallstricke epidemiologischer Berechnungen.

Ahlbohm A; Norell S: Einführung in die moderne Epidemiologie. BGA-Schriften 2/91.

AOLG: Indikatorenset der Gesundheitsberichterstattung der Länder, Bielefeld 2003.

Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit (Hrsg.): GBE Praxis 1. Gesundheitsberichterstattung für die Landkreise und kreisfreien Städte Bayern. Eine Handlungshilfe. Oberschleißheim 2004.

Beaglehole R; Bonita R; Kjellström T.: Einführung in die Epidemiologie. Verlag Hans Huber, Bern 1997.

DAE: Deutsche Arbeitsgemeinschaft Epidemiologie: Leitlinien und Empfehlungen zur Sicherung von Guter Epidemiologischer Praxis (GEP); in: Gesundheitswesen 2000; 62: 295-302.

Gordis, L: Epidemiology. W.B. Saunders, Philadelphia 1996.

Gutzwiller, F.; Jeanneret, O. (Hrsg.): Sozial- und Präventivmedizin Public Health. Verlag Hans Huber, Bern 1996.

Haisch, J; Weitkunat R; Wildner M.: Wörterbuch Public Health. Verlag Hans Huber, Bern 1999.

Hennekens CH; Buring JE.: Epidemiology in Medicine. Little, Brown and Company, Boston/Toronto 1986.

Kreienbrock, L.; Schach, S.: Epidemiologische Methoden. Gustav Fischer, Stuttgart, Jena 1997.

Kuhn, J.; Wildner, M.: Die Gesundheitsberichterstattung in Bayern: Brücke zwischen wissenschaftlicher Expertise und gesundheitspolitischer Steuerung; in: Prävention 4/2003, S. 99 ff.

Rothman KJ.: Causes. Am J Epidemiol 1976;104:587-92.

Rothman KJ.; Greenland S.: Modern Epidemiology. Second Edition. Lippincott-Raven Publishers, Philadelphia 1998.

Weiß C.: Basiswissen Medizinische Statistik. Springer, Berlin Heidelberg 1999.



91058 **Erlangen**
Eggenreuther Weg 43
Tel.: 09131/764-0



85764 **Oberschleißheim**
Veterinärstr. 2
Tel.: 089/31560-0



97082 **Würzburg**
Luitpoldstr. 1
Tel.: 0931/41993-0

www.lgl.bayern.de

**Bayerisches Landesamt für
Gesundheit und Lebensmittelsicherheit**
Eggenreuther Weg 43, 91058 Erlangen

Telefon: 09131/764-0
Telefax: 09131/764-102

Internet: www.lgl.bayern.de
E-Mail: poststelle@lgl.bayern.de

Druck: Print Com, Erlangen