

utb.

Michael Benesch
Elisabeth Steiner

Klinische Studien lesen und verstehen

2. Auflage

Leseprobe



Vorwort

In unterschiedlichen Disziplinen wie Psychologie, Soziale Arbeit, Medizin, weiteren Gesundheitsberufen oder auch in der Soziologie entstehen für statistisch bzw. epidemiologisch wenig vorgebildete Leser und Leserinnen immer wieder Schwierigkeiten, empirische Studienergebnisse zu verstehen und professionell zu beurteilen. Es gibt kaum gut lesbare, verständliche Einsteigerliteratur, welche den Studierenden sowie nicht einschlägig vorgebildeten Praktikern erläutert, wie statistische Resultate aus medizinischen und sozialwissenschaftlichen Publikationen zu verstehen, zu interpretieren und schließlich zu bewerten sind. Deshalb entschlossen wir uns, diese Lücke mit dem vorliegenden Buch „Klinische Studien lesen und verstehen“ zu schließen.

Unserem Erfolgsrezept treu bleibend, welches wir bereits bei unserem ersten gemeinsamen Lehrbuch „Der Fragebogen. Von der Forschungsidee zur SPSS-Auswertung“ (mittlerweile bereits in der dritten Auflage bei Facultas erschienen) umgesetzt haben, bedienen wir uns auch in diesem Buch einer elementaren und unkomplizierten Didaktik bei der Vermittlung komplexer Inhalte. Wir möchten vermeintlich schwierige statistische sowie epidemiologische Inhalte sehr vereinfacht und gut verständlich transportieren. Dabei wird zum Großteil auf komplexe Darstellungen von Formeln und Herleitungen verzichtet. Falls jedoch unumgänglich, gestalten wir sie gut nachvollziehbar, unnötige Zwischenschritte werden vermieden. Das Wesentliche, also das Praxisrelevante, wird in den Mittelpunkt gestellt. Somit nehmen wir die Herausforderung der Gratwanderung zwischen nötiger Tiefe und seitens der Leserinnen und Leser gewünschter Einfachheit erneut an: Unsere Erfahrung aus der universitären Lehre und zahlreichen Seminaren zeigt, dass die Vermittlung derartiger Inhalte in sehr elementarer Art und Weise große didaktische Anforderungen an Vortragende stellt.

Aus mittlerweile rund 50 Seminaren (abgehalten beispielsweise für Außendienstmitarbeiter von Pharmaunternehmen, Ärzte, Wissenschaftler verschiedener Disziplinen, Klinische Monitore, Studierende unterschiedlicher Gesundheitsberufe und Sozialer Arbeit) sind uns die Bedürfnisse unserer Zielgruppen bestens bekannt. So werden etwa regelmäßig „Checklisten zur Beurteilung medizinischer Publikationen“ gefordert, oder die Teilnehmer sprechen immer wieder bestimmte statistische Probleme an, die in Büchern

nur schwer verständlich nachzulesen sind. Den Bedürfnissen nach einfachen Erklärungen, die entweder keine oder nur sehr geringe Vorkenntnisse erfordern, soll im vorliegenden Buch nachgegangen werden. Anhand von Beispielpublikationen, die vor allem aus der Medizin, den Sozialwissenschaften und der Psychologie stammen, werden in diesem Lehrbuch die erforderlichen Kenntnisse für die Bewertung von Studien aus statistischer und epidemiologischer Sicht vermittelt. Einschlägig vorgebildete und fachlich erfahrene Personen, die explizit nicht zu unserer Zielgruppe gehören, werden uns die Vereinfachungen und Auslassungen nachsehen – der Markt ist voll von Fachbüchern, die in die Tiefe gehen, allerdings um den Preis der leichten Verständlichkeit.

Unseren Seminarteilnehmern ebenso wie unseren Studierenden sei Dank für ihr Nachfragen ausgesprochen; denn gerade diese Fragen haben uns für das Thema sensibel gemacht und die Idee für dieses Einsteigerbuch entstehen lassen. Ein besonderer Dank geht an dieser Stelle wieder an Frau MMag. Dr. Sigrid Nindl von der Facultas Verlags- und Buchhandels AG, die uns auch schon bei unseren anderen Buchprojekten unterstützt hat.

Für die überaus genaue und unterstützende Durchsicht unserer Manuskripte danken wir Frau Mag. Verena Hauser.

Wien, im Juni 2013

Michael Benesch
Elisabeth Raab-Steiner

Vorwort zur 2. Auflage

Die nun vorliegende zweite Auflage enthält einige Ergänzungen, v. a. zur Unterscheidung von Standardabweichung und Standardfehler des Mittelwerts, da dies immer wieder zu Missverständnissen bei der Beurteilung des Statistkteils medizinischer Publikationen führt. Am Grundkonzept des Buches hat sich nichts geändert, da die didaktische Zielsetzung, eine gut verständliche, leicht lesbare Einführung zur Verfügung zu stellen, ausgezeichnet angenommen wurde.

Besonderer Dank für die wie immer gute Zusammenarbeit geht an Frau MMag. Dr. Sigrid Nindl und Frau Elisabeth Klein, MA vom Facultas Verlag sowie an Frau Mag. Verena Hauser als externe Lektorin.

Waidhofen/Thaya, im Herbst 2017
Wien, im Herbst 2017

Michael Benesch
Elisabeth Steiner

Inhaltsverzeichnis

Vorwort	5
Vorwort zur 2. Auflage	7
1 Grundlagen statistischer Tests	11
1.1 Stichprobe und Grundgesamtheit	11
1.2 Signifikanz (p-Wert)	21
1.3 Die Nullhypothese und die Alternativhypothese	27
1.4 Fehler erster Art und Fehler zweiter Art	32
1.5 Das Signifikanzniveau	49
1.6 Ein- und zweiseitige Forschungshypothesen	52
1.7 Signifikanz und Relevanz, Power (Macht) und Fallzahlschätzung	56
1.8 Konfidenzintervalle	62
1.9 Standardabweichung versus Standardfehler	67
2 Medizinstatistische Kennzahlen aus der Epidemiologie ..	69
2.1 Inzidenz und Prävalenz	69
2.2 Maßzahlen für das Risiko	79
2.2.1 Odds Ratio (OR)	79
2.2.2 Relatives Risiko (RR; relative risk)	87
2.2.3 Absolute und relative Risikoreduktion (ARR, RRR)	95
2.2.4 „Number needed to treat“ (NNT; Anzahl der notwendigen Behandlungen) und „number needed to harm“	100
2.2.5 Goldstandard und diagnostische Tests	102
2.2.6 Sensitivität und Spezifität	104
2.2.7 ROC – Receiver Operator Characteristics Curve und ROC-Analysen	114
3 Die Beurteilung der Qualität von Studien	120
3.1 Hierarchie der Evidenz: Studiendesigns	122
3.2 Kriterien zur Studienbeurteilung/Checklisten	129
3.3 Eine Checkliste zur Studienbeurteilung selbst erstellen	152
Literaturverzeichnis	156
Zum Autor/Zur Autorin	159
Stichwortverzeichnis	161

1 Grundlagen statistischer Tests

Das erste Kapitel befasst sich im Wesentlichen mit dem Thema „Signifikanz“, im weiteren Sinn also mit dem p-Wert und allem, was sich um diesen dreht. Wenn eine Publikation betrachtet wird, steht sehr oft die Frage im Mittelpunkt, ob das Ergebnis signifikant ist oder nicht. Dabei ist der viel beachtete p-Wert nur einer von mehreren relevanten Faktoren, und die Kenntnis des p-Werts alleine reicht bei Weitem nicht aus, um die Bedeutung einer Studie abschätzen zu können. Deshalb ist es wichtig, einige statistische Prinzipien zu kennen, die mit dem „statistischen Test“ verbunden sind. Diesem Ziel dient das folgende Kapitel.

1.1 Stichprobe und Grundgesamtheit

Um welches Problem geht es in diesem Unterkapitel?

In Österreich leiden etwa 11 % der Erwachsenen zwischen 18 und 65 Jahren an Adipositas. Möchte man wissen, ob eine neue Therapie bei diesen rund 11 % wirksam ist, hat man folgendes Problem zu lösen: Man wird nur einen Teil dieser Patienten untersuchen können, sagen wir: $n = 300$ Personen („ n “ steht für „number“ und bezeichnet die Stichprobengröße). Die gesamten 11 % der adipösen Patienten bilden die Grundgesamtheit oder Population (diese Population besteht also aus ein paar Hunderttausend Menschen) und die für die Untersuchung ausgewählten $n = 300$ Personen bilden die Stichprobe (das Sample). Das Problem dabei: Wie können wir etwas über die Population aussagen, wenn wir nur $n = 300$ Personen untersucht haben?

Wenn Sie dieses Unterkapitel gelesen haben, können Sie folgende Fragen beantworten:

Was ist der Unterschied zwischen Stichprobe (Sample) und Grundgesamtheit (Population)?

Warum ist es notwendig, Zufallsstichproben (engl.: random samples) zu ziehen?

Was kann passieren, wenn der Experimentator-Erwartungs-Effekt wirkt?

Welcher Unterschied besteht zwischen beschreibender (deskriptiver) und schließender (Inferenz-)Statistik?

Schlüsselbegriffe

Zufallsstichprobe, Population, Grundgesamtheit, Deskriptivstatistik, Inferenzstatistik

Stellen Sie sich bitte Folgendes vor: Ein Landwirt besitzt 100 Milchkühe. Da in der Gegend um seinen Hof eine Parasitose grassiert, möchte er wissen, wie viele seiner Tiere erkrankt sind. Der Serumtest ist sehr teuer, also kann er vom Tierarzt nur 30 Rinder testen lassen. In diesem Fall ist die Gesamtheit seiner Kühe die Population ($N = 100$), und wenn der Veterinär $n = 30$ Kühe ausgewählt hat, bilden diese die Stichprobe. (Ein „großes N “ bezeichnet meist die Population [Grundgesamtheit], ein „kleines n “ die Stichprobe.)



Die 30 Stichproben-Rinder müssen per Zufall ausgewählt werden, am besten mithilfe eines Zufallsgenerators. Denn wenn der Landwirt mit verbundenen Augen auf die Weide geht und so per Zufall 30 Stück herausgreift, könnte Folgendes passieren: Die kranken Tiere gruppieren sich, zum Beispiel weil die Parasitose sie träge macht, in einem Teil des Weidegrundes, sodass der Tierarzt vor allem die kranken Rinder testet. In diesem Fall wäre der Anteil kranker Tiere in der Stichprobe überhöht. Dieser Fehler (man spricht von einem Selektionsbias, von engl. bias: systematischer Fehler) kann ausgeschlossen werden, wenn jede Kuh eine Nummer erhält und per Zufallsgenerator gezogen wird. Ähnliches kann natürlich auch passieren, wenn im Rahmen einer medizinischen Studie die Probanden (allgemein: die Elemente) nicht zufällig, das heißt randomisiert (von engl. random: zufällig), ausgewählt werden.

Bereits in den 1950er Jahren wurde im Rahmen der Salk-Polio-Studie das Konzept der Randomisierung erstmals in großem Stil in der medizinischen Forschung eingesetzt (Schumacher & Schulgen, 2002). Man wollte untersuchen, ob der von Jonas Salk entwickelte Impfstoff die Zahl der Polio-Neuerkrankungen reduzieren kann. Unter anderem wurde daran gedacht, Eltern die freiwillige Teilnahme ihrer Kinder anzubieten, um so die Inzidenz (Neuerkrankungsrate) in der Gruppe der Geimpften mit jener

der Nichtgeimpften zu vergleichen. Allerdings hätte es hier passieren können, dass die Zustimmung seitens der Eltern stark vom sozioökonomischen Status abhängt, der seinerseits vermutlich wiederum das Auftreten der Erkrankung beeinflussen könnte. Es wurde schließlich folgender Ausweg gewählt: Jene Kinder, die an der Studie teilnehmen durften, wurden randomisiert der Verum- und Placebogruppe zugeteilt (als Verum bezeichnet man das „wahre“ Medikament im Gegensatz zum Placebo, das – abgesehen vom Placeboeffekt – wirkungslos ist). Auf diese Weise konnten derartige Effekte praktisch ausgeschlossen werden.

Der Tierarzt zieht also per Zufallsgenerator aus den 100 Rindern eine Stichprobe von 30 Tieren. Wie man sich leicht denken kann, gibt es sehr viele verschiedene Möglichkeiten, 30 aus 100 Tieren auszuwählen. In Abbildung 1.1 sind fünf mögliche Stichproben dargestellt.

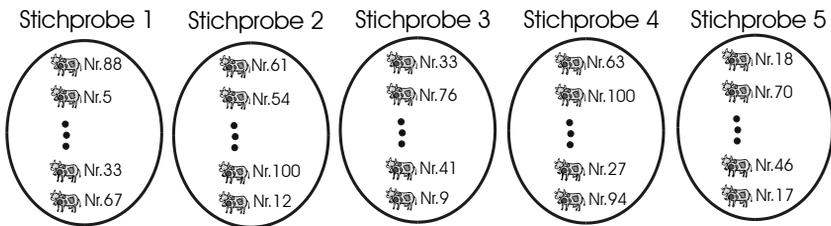


Abb. 1.1: Fünf mögliche Stichproben

Fünf verschiedene Stichproben werden sich in der Zusammensetzung der aufgenommenen Tiere unterscheiden (es sind ja Zufallsstichproben). Nun wird in jeder dieser fünf (hypothetischen) Stichproben der Anteil kranker Kühe ermittelt (siehe Tab. 1.1):

Tab. 1.1: Resultate aus fünf möglichen Stichproben

Stichprobe	Anteil kranker Kühe
1	9 %
2	9 %
3	12 %
4	8 %
5	10 %

Es ist anzunehmen, dass der Anteil kranker Rinder nicht in allen fünf Stichproben gleich sein wird – manchmal werden mehr, manchmal weniger kranke Tiere enthalten sein, da es sich ja um Zufallsziehungen handelt (die Elemente [= Rinder] wurden per Zufall ausgewählt). Man spricht von Stichprobenergebnissen, die je nach Stichprobe – aufgrund der Zufallsauswahl – zufällig schwanken. Diese Stichprobenergebnisse werden vermutlich vom Anteil kranker Rinder in der Population ($N = 100$) abweichen. Natürlich kann es sein, dass ein Stichprobenresultat dem Anteil kranker Tiere in der Population entspricht, aber da der Tierarzt nicht die gesamte Rinderpopulation untersucht hat, weiß er darüber nicht Bescheid.

In den meisten Fällen ist man nicht an Stichprobenergebnissen interessiert, sondern daran, wie es in der Population aussieht. Zum Beispiel: Wie hoch ist der durchschnittliche Blutdruck bei allen Personen, die an einer bestimmten Krankheit leiden? Da nicht alle Kranken untersucht werden können, wird also von der Stichprobe, und zwar von einer der vielen möglichen Stichproben, auf die Grundgesamtheit (Population) geschlossen. Dieses Schließen von der Stichprobe auf die Population ist stets mit Unsicherheiten behaftet, da man die Population i. d. R. ja nicht kennt. Bei Wahlprognosen äußert sich dies darin, dass ein Schwankungsbereich angegeben wird.

Unterschied Inferenzstatistik – Deskriptivstatistik

Wollten wir einfach eine Stichprobe beschreiben und nichts über die Population wissen, bräuchten wir die Wahrscheinlichkeiten nicht zu bemühen. Wenn wir nur die Stichprobe von $n = 15$ Rindern beschreiben möchten, indem wir fragen: „Wie hoch ist der Anteil kranker Rinder in dieser Stichprobe?“, dann ermitteln wir diesen Anteil und belassen es dabei. (Man spricht von Deskriptivstatistik, von lat. *describere*: beschreiben.) Möchten wir allerdings aufgrund dieser Stichprobe auf den Anteil kranker Rinder in der Population schließen, dann geht dies nur noch mit mehr oder minder unsicheren Wahrscheinlichkeitsaussagen. (Man spricht von Inferenzstatistik, von lat. *infero*: folgern, schließen.) Also: Das Schließen von der Stichprobe auf die Population wird damit erkauft, dass die Aussagen unsicher sind, sie können nur mit einer gewissen Wahrscheinlichkeit getroffen werden. Der p-Wert liefert uns Informationen darüber, wie wahrscheinlich ein bestimmtes Stichprobenresultat unter bestimmten Vorannahmen ist (dazu später mehr).

Die Deskriptivstatistik beschreibt eine Stichprobe, die Inferenzstatistik trifft Aussagen über die Population basierend auf der Stichprobe. Inferenzstatistische Aussagen sind mehr oder minder unsicher und können nur mit einer bestimmten Wahrscheinlichkeit getroffen werden, da von einer der vielen möglichen Stichproben auf die i. d. R. unbekannte Grundgesamtheit geschlossen wird.

Aus einer Studie

In dieser Studie wurde untersucht, ob Maßnahmen zur Hygieneschulung in bestimmten ländlichen Gebieten Zimbabwes wirken. In sogenannten Community Health Clubs vermitteln Trainer (Environmental Health Technicians: EHTs) hygienerelevante Aspekte an die Bevölkerung. Aus insgesamt 297 Clubs in den Provinzen Tsholotsho und Makoni wurden 25 ausgewählt.

“A systematic sample of 25 clubs was taken from the NGO’s register of Health Clubs in each district to include clubs from each ward ... The sample of clubs in each district was chosen to include at least one club from each ward, and so prevent bias arising from different EHTs’ methods.”

Waterkeyn, J. & Cairncross, S. (2005). Creating demand for sanitation and hygiene through Community Health Clubs: A cost-effective intervention in two districts in Zimbabwe. *Social Science & Medicine*, 61, 1958–1970.

Die in die Studie eingeschlossenen 25 Clubs stellen somit die Stichprobe (Sample) aus der Population von 297 Clubs dar. Um mögliche Verzerrungen zu vermeiden, die aus unterschiedlichen Arbeitsweisen der Environmental Health Technicians (EHTs) resultieren könnten, wurde darauf geachtet, zumindest einen Club jedes Clubleiters mit einzubeziehen. Das optimale Vorgehen wäre natürlich, alle 297 Clubs der beiden Provinzen zu untersuchen. Aus organisatorischen und Kostengründen ist ein derartiges optimales Vorgehen allerdings meist nicht möglich, weshalb Stichproben gezogen werden. Aus diesen und weiteren – etwa ethischen – Überlegungen werden beispielsweise bei Medikamentenstudien auch nicht alle Patienten (diese bilden die Population) in eine Studie aufgenommen, sondern nur Stichproben.

2.2.2 Relatives Risiko (RR; relative risk)

Um welches Problem geht es in diesem Unterkapitel?

Bei der frühzeitigen Diagnose einer HIV-Infektion kann eine vorübergehende antiretrovirale Therapie sinnvoll sein. Dieser Therapieerfolg konnte durch eine randomisierte klinische Studie nachgewiesen werden und wurde im *New England Journal of Medicine* (Le et al. 2013, 207–217) publiziert. Es wurde bei der Behandlung die Beobachtung gemacht, dass es durch die Therapie zu einer beschleunigten Erholung der Immunparameter kam.

Um den Wirkungsunterschied – also das Maß für die Stärke des Zusammenhangs zwischen Therapie und Zielereignis – zu finden, berechnet man das relative Risiko.

In der Medizin werden häufig Interventionsstudien durchgeführt, die den Nachweis der Wirksamkeit einer (meist neuen) Interventionsmaßnahme bzw. eines neuen Medikaments erbringen sollen. Ist ein neues Medikament oder eine neue Operationstechnik erfolgversprechend? Entsteht beim Einsatz dieser neuen Technik oder dieses neuen Medikaments ein Nutzen für die betroffenen Personen, oder kommen diese möglicherweise auch zu Schaden? Wie kann der Effekt der Therapie (der Wirkung) erfasst werden?

Wenn Sie dieses Unterkapitel gelesen haben, können Sie folgende Fragen beantworten:

Was versteht man unter einem relativen Risiko?

Wann ist ein relatives Risiko zu berechnen?

Wie lassen sich nach einer Berechnung die erzielten Ergebnisse richtig interpretieren?

Schlüsselbegriffe

Risiko, relatives Risiko, relative risk, risk ratio, Risikodifferenz, absolute Risikoreduktion, relative Risikoreduktion, Number needed to treat

Ausgehend vom Nutzen bzw. Nachweis einer Intervention bzw. eines Medikaments kann entweder ein Vergleich mit einer gegenwärtig eingesetzten Standardmethode oder mit der Anwendung eines Placebos erfolgen. Die teilnehmenden Personen werden den Gruppen zufällig (randomisiert, siehe Kap. 1.1) zugeteilt. Ziel ist die Schätzung des Effekts durch die Intervention in den beiden Gruppen – bei dieser Schätzung des Effekts spricht man auch von „Risiko“.

Bei der Ermittlung des relativen Risikos betrachtet man jeweils den Anteil der Gruppe der exponierten Personen an der Gesamtzahl der Patienten in den beiden entsprechenden Therapiegruppen, das heißt, die Wahrscheinlichkeiten werden betrachtet.

„Das Risiko (griech. Klippe, Gefahr) kennzeichnet in der Regel den erwarteten **Ausgang eines unerwünschten Ergebnisses**. Aus mathematischer Sicht ist das Risiko ein Produkt aus der Wahrscheinlichkeit mit den Konsequenzen (Kosten) aus den möglichen Ausgängen eines Ereignisses und somit nicht dimensionslos.“ (Hedderich/Sachs 2012, S. 140) Der Unterschied zur Odds Ratio ist, dass sich das relative Risiko nicht auf Quoten, sondern auf Wahrscheinlichkeiten bezieht!



Zur Wiederholung: Die Odds Ratio ist ein Quotient zweier Odds, z. B. Chance für das Auftreten des Zielereignisses unter Therapie 1 dividiert durch die Chance für das Auftreten des Zielereignisses unter Therapie 2. Das relative Risiko einer Therapie 1 zu einer Therapie 2 ist der Quotient der Wahrscheinlichkeit für das Eintreten des Zielereignisses unter den beiden Therapien. Mit dem relativen Risiko (relative risk) kann ausgedrückt werden, um welchen Faktor sich ein Risiko zum Beispiel für eine Erkrankung in zwei Gruppen unterscheiden lässt. Es ist eine Begrifflichkeit der deskriptiven Statistik.

„Das Relative Risiko (RR) wird berechnet, indem das Risiko in der gegenüber dem Risikofaktor exponierten Gruppe durch das Risiko in der nicht exponierten Gruppe dividiert wird.“ (Rossoth 2007, S. 45)

$$\text{Relatives Risiko} = \frac{\text{Anteil in der exponierten Gruppe}}{\text{Anteil in der Kontrollgruppe}}$$

Es gibt an, wie hoch die Wahrscheinlichkeit ist, dass ein Merkmal (von zwei Alternativmerkmalen) für eine Gruppe (von zwei Gruppen) vorliegt.

Diese Wahrscheinlichkeiten werden in diesem Kontext üblicherweise als Risiken bezeichnet. Davon ausgehend spricht man daher zum Beispiel vom Risiko für einen Therapieerfolg unter der experimentellen bzw. Standard- oder

Placebobehandlung. Um einen Vergleich anstellen zu können, werden gebräuchliche Maßzahlen zum Vergleich der beiden Gruppen aus den Wahrscheinlichkeiten abgeleitet, nämlich ein Quotient. Dieser Quotient gibt wiederum die Möglichkeit, etwas über die Stärke der Assoziation zwischen der Exposition und der Krankheit auszusagen, und ist daher gut zur Klärung von Krankheitsursachen geeignet.

So wird zum Beispiel Hepatitis B mit dem Risiko für Leberzellkarzinom assoziiert. Nach neun Jahren der Infektion mit Hepatitis B liegt das relative Risiko für Leberzellkarzinom bei 98,4, das heißt, Patienten mit Hepatitis B erkranken 98-mal häufiger an Leberzellkarzinom als Patienten, die keine Hepatitis haben (vgl. Rosboth 2007, S. 46). Hier liegt ein negativer Effekt vor. Für positive Ereignisse ist der Begriff des Risikos eher unpassend und eigentlich nicht zutreffend, dennoch hat sich die Begrifflichkeit in der medizinischen Literatur durchgesetzt. Vorzuziehen wäre der Begriff Wahrscheinlichkeit.

Die Basis für die Herleitung bzw. die **Berechnung des relativen Risikos** stellt wiederum eine Vierfeldertafel dar (siehe Tab. 2.5). In epidemiologischen (medizinischen) Studien bzw. Statistiken sind folgende Angaben sehr häufig und sollen als Grundschemata für die Erklärung des relativen Risikos dienen:

Gefüllt wird diese Tabelle mit den Häufigkeiten der Personen, die

- krank mit Risikofaktor (a),
- nicht krank mit Risikofaktor (b),
- krank ohne Risikofaktor (c) und
- nicht krank ohne Risikofaktor (d) sind.

Tab. 2.5: Grundschemata einer Vierfeldertafel für epidemiologische Studien

	Anzahl der erkrankten Personen	Anzahl der nicht erkrankten Personen	Gesamt
Anzahl der Personen mit Risikofaktor	a	b	a + b
Anzahl der Personen ohne Risikofaktor	c	d	c + d
Gesamt	a + c	b + d	

Im ersten Schritt wird „Krankheit mit Risikofaktor“, also $\frac{a}{a+b}$ berechnet.

Im zweiten Schritt „Krankheit ohne Risikofaktor“, also $\frac{c}{c+d}$.

Daraus wird der Quotient gebildet – zur Berechnung des relativen Risikos (RR) dient folgende Formel:

$$RR = \frac{\frac{a}{a+b}}{\frac{c}{c+d}}$$

$$RR = \frac{\text{Erkrankung/Personen mit Risikofaktor}}{\text{Erkrankung/Personen ohne Risikofaktor}}$$

Ein Beispiel:

Es wurde eine Metaanalyse (Edwards et al. 2004) zur Wirksamkeit von selektiven Zyklooxygenase-2-Hemmern im Vergleich zu Placebo bei akuten postoperativen Zahnschmerzen durchgeführt. Im Rahmen dieser Studie wurde auch für den nichtselektiven Zyklooxygenase-2-Hemmer Diclofenac (50 mg) eine Einzelstudie durchgeführt, aus der Daten vorliegen, die in Tabelle 2.6 dargestellt sind. In dieser finden sich die Angaben von Personen, die zur Reduktion der postoperativen Zahnschmerzen um mindestens 50% innerhalb von sechs Stunden einerseits Diclofenac 50 mg oder ein Placebo verabreicht bekommen haben.

Mit dieser Studie sollte ein Therapieerfolg als Zielgröße untersucht werden.

Tab. 2.6: *Wirksamkeit von Diclofenac im Vergleich zu Placebo bei akuten postoperativen Zahnschmerzen*

	Reduktion der Zahnschmerzen um mind. 50 % (sechs Stunden)		Σ
	ja	nein	
Diclofenac 50 mg	a 32	b 89	121
Placebo	c 8	d 55	63
Σ	40	144	184

Für die Verabreichung von 50 mg des Medikaments Diclofenac ergibt sich bezüglich der Reduktion der Zahnschmerzen um mindestens 50 % innerhalb der ersten sechs Stunden nach der Einnahme:

- Wahrscheinlichkeit „Reduktion der Zahnschmerzen mit Diclofenac“: $32/121 = 0,26 = 26\%$
- Wahrscheinlichkeit „Reduktion der Zahnschmerzen mit Placebo“: $8/63 = 0,127 = 12,7\%$

$$RR = \frac{0,26}{0,127} = 2,08$$

Das „Risiko“ (bzw. die Erfolgswahrscheinlichkeit) für einen Therapieerfolg ist in der Diclofenac-Gruppe zweimal so hoch wie in der Placebogruppe, allein die Beobachtung der Prozentwerte verdeutlicht das. Das „Erfolgsrisiko“ bei Exponierten ist zweimal so groß wie bei Nichtexponierten.

Relatives Risiko: Maßzahl zur Darstellung des Wirkungsunterschiedes. Es ist ein Maß für die Stärke des Zusammenhangs zwischen der Therapie und dem Zielereignis.

Fassen wir die Eigenschaften des relativen Risikos zusammen:

- Das relative Risiko nimmt Werte zwischen 0 und $+\infty$ an.
- Ein Wert von 1 bedeutet, dass das Risiko in beiden Gruppen gleich ist. Es besteht dementsprechend kein Anhaltspunkt für einen Zusammenhang zwischen der untersuchten Erkrankung und dem Risikofaktor.
- > 1 : Liegt das relative Risiko über 1, ist das ein Hinweis auf einen möglichen positiven Zusammenhang zwischen einem Risikofaktor wie beispielsweise Rauchen und einer Erkrankung.
- < 1 : Liegt das relative Risiko unter 1, hat die Exposition eine protektive (schützende) Wirkung, wie dies beispielsweise bei Impfungen der Fall ist.

Aus einer Studie

In einem anerkannten amerikanischen Journal für Geburtshilfe und Gynäkologie (BJOG: An International Journal of Obstetrics and Gynaecology) wurde im Mai 2013 in einer klinischen Studie der Zusammenhang zwi-

Studierende wie Lehrende stehen häufig vor der Herausforderung, Studien lesen und deren Ergebnisse bewerten zu müssen. Dieses verständliche Lehrbuch vermittelt die Grundlagen dafür und erscheint nun in der zweiten, aktualisierten und überarbeiteten Auflage. Statistisch bzw. epidemiologisch nicht oder kaum vorgebildete Leser lernen an Beispielen der Medizin, Sozialwissenschaften oder Psychologie, Studienergebnisse zu beurteilen und zu verstehen.

Zentral sind das „Schlüsselthema Signifikanz“ und die damit verbundenen Begriffe (z. B. Alpha-Fehler, Power Analyse, Signifikanzniveau, p-Wert).

Weiters werden häufige medizinstatistische Kennzahlen erklärt und können mithilfe des Buches korrekt interpretiert werden. Checklisten bieten zudem praktische Hilfen bei der systematischen Beurteilung von Publikationen.

Dies ist ein utb-Band aus dem Verlag facultas.
utb ist eine Kooperation von Verlagen mit einem gemeinsamen Ziel: Lehrbücher und Lernmedien für das erfolgreiche Studium zu veröffentlichen.

ISBN 978-3-8252-4896-3



9 783825 248963



QR-Code für mehr Infos und
Bewertungen zu diesem Titel

utb-shop.de