

Interpretation epidemiologischer Daten

Im dritten Artikel dieser Serie gibt Mona Okasha eine schrittweise Anleitung zum Verständnis einer epidemiologischen Studie

Der Artikel des letzten Monats bezog sich auf bedeutende Aspekte zum epidemiologischen Studiendesign [1]. Der Fokus in diesem Artikel liegt darauf, wie die Ergebnisse einer Studie interpretiert werden. Dabei soll dieser Leitfaden gleichermaßen anwendbar sein sowohl für Ihre eigenen Studien als auch für publizierte Artikel in Fachzeitschriften. Wenn Sie die hier aufgezeigten Strukturen verwenden, werden Sie in der Lage sein, auch eine epidemiologische Studie zu bewerten, die Ihnen im Examen vorgelegt wird.

Zwischen dem Studiendesign und der Interpretation der Ergebnisse liegt ein weites Feld – die statistische Analyse. Statistiken werden eingesetzt, um mit den Daten umzugehen, die die Ergebnisse bilden. Dies ist ein großes eigenes Spezialgebiet, für welches ich auf das Buch von Kirkwood (s. weiterführende Literatur) für einen klaren Überblick über dieses Gebiet verweisen möchte.

Die Sprache der Epidemiologie

Wie in anderen medizinischen Disziplinen, so hat auch die Epidemiologie ihr eigenes Vokabular. Ich werde in diesem Artikel Begriffe verwenden, die Ihnen eventuell unvertraut sind. Erläuterungen hierzu werden im Glossar gegeben. Die Epidemiologie wird eingesetzt, um Zusammenhänge zwischen Ereignissen und Ergebnissen zu beschreiben. In einer Studie zu Sexualität und Tod, in der untersucht wurde, ob Männer mittleren Alters, die häufige Orgasmen haben, eher sterben als ihre weniger sexuell aktiven Altersgenossen, ist das Ereignis die Frequenz der Orgasmen und das Ergebnis ist der Tod [2].

Die Ergebnisse wörtlich nehmen

Die Ergebnisse epidemiologischer Studien werden als Vergleich der Befunde von zwei oder mehreren Gruppen ausgedrückt. In der Sexualitätsstudie wollen wir wissen, ob die Todesrate zwischen Gruppen von Männern mit unterschiedlich häufigen Orgasmen vergleichbar ist. Wir ermitteln die Rate bei Männern mit unregelmäßigen und bei jenen mit häufigen Orgasmen. Das Verhältnis der ersten gegenüber der zweiten Gruppe ergibt den Verhältniswert oder die Korrelation (r). Wenn die Raten in den beiden Gruppen gleich sind, beträgt die Korrelation 1. Eine Korrelation von $r < 1$ bedeutet eine geringere Rate unter den Männern mit wenigen Orgasmen verglichen mit den Männern mit häufigen Orgasmen. Eine Korrelation von $r > 1$ bedeutet, dass die Rate unter den Männern mit einigen Orgasmen geringer ist als diejenige der Männer ohne Orgasmen. In der Sexualitätsstudie wurde das Verhältnis der Ungleichheit (Odds Ratio, OR) verwendet, um sich der Korrelation zu nähern. Der ermittelte Wert von 2,0 zeigt an, dass Männer mit seltenen Orgasmen doppelt so häufig in einem festgelegten Zeitraum sterben wie Männer mit häufigen Orgasmen.

Glauben wir an diese Ergebnisse?

Statistische Werte wie Korrelationen werden eingesetzt, um Zusammenhänge zu beschreiben. Wir müssen dann entscheiden, wie viel Geltung wir diesen Ergebnissen zuschreiben. Wir bewerten die Ergebnisse unter 5 Gesichtspunkten (siehe Kasten). Die Reihenfolge, in der

sie aufgeführt sind, ist wichtig, weil auf jeder Stufe, auf der die Studie nicht einwandfrei bewertet wird, muss dies für alle weiteren Betrachtungen bedacht werden.

Kriterien zur Bewertung epidemiologischer Ergebnisse

- Zufallsstreuung
- Verzerrungen – Abweichung von der Wahrheit; Selektionsfehler, wenn Teilnehmer nicht per Zufall ausgewählt wurden; Informationstendenzen, wenn die Genauigkeit der Daten sich hinsichtlich der Ereignisse oder der Ergebnisse unterscheiden
- Störfaktoren – alternative Erklärung
- Umgekehrte Kausalität – vielleicht verursacht das Ergebnis das Ereignis und nicht umgekehrt
- Kausaler Zusammenhang – dies ist eine überzeugende Annahme, Vorsicht

Zufall

Die Statistik hilft uns, zu entscheiden, ob unsere Ergebnisse möglicherweise Zufallsbefunde sind. Es gibt keinen definitiven Schwellenwert, ab dem man sagen könnte, es handelt sich um Zufall oder nicht um Zufall. Dies wurde früher durch die Verwendung von Wahrscheinlichkeitswerten getan. Dieses Verfahren wird mittlerweile in der Epidemiologie nicht mehr als angemessen angesehen [3]. Um zu glauben, ob ein Ergebnis auf Zufall oder nicht zurückzuführen ist, ist eine qualitative Entscheidung, die unterstützt wird durch die Berechnung eines Vertrauensintervalls (Confidence Interval CI). Die OR ist nur ein Schätzwert. Wenn man die gleiche Studie nochmals durchgeführt

hätte, könnte man möglicherweise leichte Ergebnisabweichungen erhalten. Das CI vermittelt uns einen Bereich, in dem wir relativ sicher sein können, dass sich darin die „wahre“ OR befindet. In der Sexualitätsstudie reicht das 95%ige Vertrauensintervall von 1,1 bis 3,5. Frei interpretiert können wir zu 95 % sicher sein, dass die wahre OR zwischen diesen beiden Werten liegt. Zur Erinnerung, wenn die Werte in beiden Gruppen gleich sind, ist die OR 1,0.

Fehler

Wenn wir froh darüber sind, dass unsere Ergebnisse nicht einfach eine Zufallsangelegenheit sind, bewerten wir danach, ob sie von Fehlern beeinflusst und unausgewogen oder schlichtweg falsch sein könnten. Unausgewogenheit ist grundsätzlich ein Problem des Studiendesigns oder der Datenerhebung und sollte möglichst von Anbeginn der Studie an vermieden werden. Wenn Sie die Ergebnisse eines anderen interpretieren, rechnen Sie stets damit, dass Unausgewogenheiten an allen Stellen der Studie aufgetreten sein können. Es gibt zwei Hauptfehlerquellen, den Selektionsfehler und die Informationsfehler.

Selektionsfehler

Der Selektionsbias bezieht sich auf einen Fehler bei der Auswahl von Menschen für eine Studie. Ein Selektionsfehler entsteht, wenn die Gründe für die Entscheidung zur Aufnahme oder zum Ausschluss von der Studie in Beziehung zum Ereignis oder zum interessierenden Ergebnis steht. Der Selektionsbias ist ein bedeutendes Thema in Fall-Kontroll-Studien. Denken Sie an eine Studie zu Alkoholkonsum und Motorradunfällen. Wir könnten die Fälle in einer Unfall- oder Notaufnahme erheben. Die Kontrollfälle desselben Krankenhauses wären ärmer, würden mehr rauchen und mehr Alkohol konsumieren (allgemeine Charakteristika einer stationären Population) als die Allgemeinbevölkerung. Weil die Entscheidung zur Auswahl der Kontrollfälle (stationäre Patienten) in Beziehung steht zum Ereignis (Alkoholkonsum), können unsere Ergebnisse fehlerhaft sein. Ist die Studie einmal begonnen,

können wir nicht viel tun, um das Problem zu beheben.

In Kohortenstudien ist der Selektionsbias ein Thema, wenn Teilnehmer aus solchen Gründen aus der Studie ausscheiden, die mit dem Ereignis oder dem Ergebnis der Studie zu tun haben. In der Sexualitätsstudie sind Männer, die nicht teilnahmen, eher älter und eher in Berufen mit körperlicher Tätigkeit als diejenigen Männer, die die Fragen zu den Orgasmen beantwortet haben. Die aus der Studie ausgeschlossenen Männer könnten eher häufige Orgasmen haben oder eher versterben. Das Verhältnis zwischen den Mortalitätsraten, das in dieser Studie ermittelt wurde, ist wahrscheinlich kleiner als das Verhältnis, das gefunden worden wäre, wenn alle Männer einbezogen wären. Der Selektionsfehler kann sich auch in die gegenteilige Richtung auswirken und die Ergebnisse über den wahren Wert erhöhen.

Informationsfehler

Der Informationsfehler entsteht, wenn wir Informationen fehlerhaft erheben. Er ist auch bekannt als differenzieller Messfehler, wenn wir zwischen Fällen und Nicht-Fällen bzw. Kontrollen unterschiedlich messen. In Fall-Kontroll-Studien kommt es häufig zu Informationsfehlern, wenn Fälle und Kontrollen sich unterschiedlich an bestimmte Gegebenheiten erinnern oder Dinge falsch wiedergeben. Menschen, die erkrankt sind, sagen eher als Kontrollfälle, dass sie dünner waren, weniger rauchten oder dergleichen sozial erwünschte Eigenschaften. Dieses Phänomen ist auch bekannt als der Wiedererkennungsbias und ist eine bedeutende Fehlerquelle in Fall-Kontroll-Studien. In Kohortenstudien tritt der Informationsfehler vor allem bei der Messung der Ergebnisse auf. Zum Beispiel werden in Behandlungsstudien diejenigen Teilnehmer genauer beobachtet, als „gesunde“ Teilnehmer. Die Ergebnisermittlung (Erfassung der Fälle) kann zwischen der exponierten und der nicht exponierten Gruppe unterschiedlich sein. In der Sexualitätsstudie gibt es keinen Grund zu denken, dass die Autoren eher Todesfälle von Männern mit häufigen oder seltenen Orgasmen gefunden haben, weshalb es unwahrscheinlich ist, dass der Informationsfehler ein Problem darstellt.

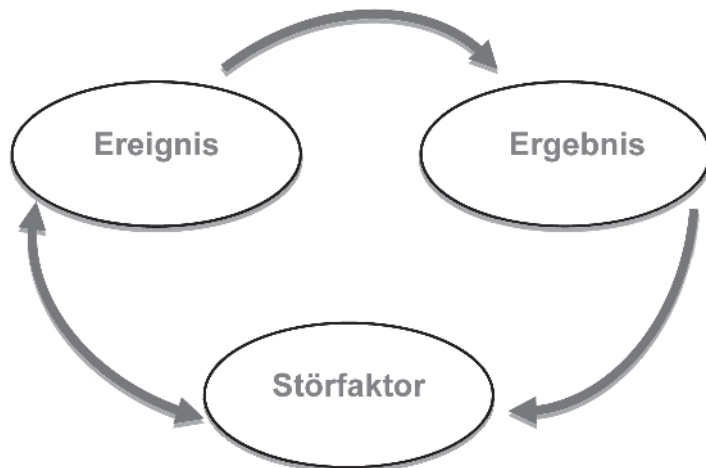
Störfaktoren

Wir sind nun soweit zu denken, dass unsere Ergebnisse keine zufälligen Beobachtungen sind und nicht durch mögliche Fehlerquellen oder ein Übersehen unsererseits entstanden sind. Bevor wir uns auf die Schlussfolgerung stürzen, dass das Ereignis das Ergebnis verursacht oder verhindert – zum Beispiel Sex verzögert das Sterben – müssen wir alternative Erklärungen für das, was wir gesehen haben, erwägen. Wenn es eine andere mögliche Erklärung gibt, sagen wir, dass der Zusammenhang zwischen Ereignis und Ergebnis von irgendetwas gestört wird. Dieses „Irgendetwas“ wird Störfaktor, Konfounder oder konfundierende Variable genannt. Um ein Störfaktor zu sein, muss dieser in Beziehung zum Ereignis stehen und ein Risikofaktor für das Ergebnis sein (siehe Graphik).

Der allgegenwärtige Störfaktor ist das Alter. Die meisten Krankheiten werden mit dem Alter häufiger (demnach ist der Störfaktor ein Risikofaktor für das Ergebnis). Viele Ereignisse besonders verhaltensbezogene Faktoren, sind ebenfalls altersabhängig. Im Gegensatz zu den Fehlern gibt es Möglichkeiten, solchen Störeinflüssen zuvorzukommen. Mit der Statistik können wir Ergebnisse um Störfaktoren adjustieren oder kontrollieren. Diese Adjustierung berücksichtigt den Störfaktor und entfernt seinen Einfluss. Ein Vergleich der rohen (nicht adjustierten) Ergebnisse mit den angeglichenen Ergebnissen zeigt auf, ob störende Einflüsse vorliegen. In der Sexualitätsstudie veränderte sich die Odds Ratio von 2,0 auf 1,9, nachdem bezüglich Alter, sozialer Schicht, systolischem Blutdruck, Rauchen und Koronarer Herzkrankheit adjustiert wurde, weshalb keine Störfaktoren anzunehmen waren. Wir können allerdings nur nach solchen Faktoren adjustieren, die auch bekannt sind und weitere Störfaktoren können verbleiben, wenn die adjustierten Ergebnisse weiterhin unter dem störenden Einfluss anderer Variablen leiden.

Umgekehrte Kausalität

Jetzt haben wir das Stadium erreicht, in dem wir denken, dass die Ergebnisse, die wir haben, nicht durch Zufall, Fehler oder Störfaktoren erklärt werden können. Bevor erwogen wird, ob das Ereignis



nis möglicherweise das Ergebnis verursacht, müssen wir berücksichtigen, ob das Ergebnis möglicherweise das Ereignis verursacht haben kann. Dies wird umgekehrte Kausalität genannt. In der Studie zu Sexualität und Tod ist dies eine wahrscheinliche Erklärung der Ergebnisse. Männer, die zu Beginn der Studie an einer chronischen Krankheit litten, hatten ein größeres Risiko während des Untersuchungszeitraums zu sterben und sind wahrscheinlich weniger sexuell aktiv gewesen. Die Krankheit (Ergebnis) hat Individuen veranlasst ihr Verhalten (Ereignis) zu ändern. Dies hat eine große Bedeutung in Fall-Kontroll-Studien und Kohortenstudien mit kurzen Beobachtungszeiträumen.

Kausalität

Es ist ein großer Sprung von der Beobachtung eines Zusammenhangs bis zur Überzeugung, dass das Ereignis das Ergebnis verursacht. Kriterien, die erfüllt sein sollten, bevor Kausalität angenommen werden kann, wurden von Austin Bradford Hill im Jahr 1965 dargelegt [4]. In Kürze, diese Kriterien verlangen konsistente Beweise, die zusammengekommen sind. Das Ereignis muss eindeutig dem Ergebnis zuvorkommen, und es sollte biologische oder andere rationale Erklärungen der Ergebnisse geben. Es ist wichtig sich daran zu erinnern, dass eine einzelne epidemiologische Studie nie-

mals ausreicht, um eine Kausalität zu begründen.

Ist Sex also gut für uns?

Jetzt sollten Sie in der Lage sein, Davey Smith's et al. Artikel kritisch zu bewerten und eine informierte Entscheidung zu treffen, ob Sie glauben, dass ein Mangel an regelmäßigen Orgasmen Ihr Mortalitätsrisiko erhöht. Bedenken Sie, dass es niemals eine richtige oder falsche Antwort gibt. Das Bewerten der Epidemiologie besteht darin, die Evidenz abzuwägen und zu einer Schlussfolgerung zu kommen, die in Übereinstimmung steht sowohl mit den experimentellen als auch den beobachteten Ergebnissen.

Anmerkung

Teil 3 einer Serie aus dem Student BMJ, Volume 9, September 2001, S. 324–325, übersetzt von Dr. Andreas Loh (Freiburg). Abdruck der Übersetzung mit freundlicher Genehmigung der BMJ Publishing Group.

Korrespondenzadresse:

Mona Okasha
Epidemiologin, Universität Bristol
E-Mail: Mona.Okasha@bristol.ac.uk

Glossar

- **Adjustierung:** Ergebnisse verändern, weil Störfaktoren Zusammenhänge beeinflussen.
- **Assoziation:** Die Beziehung zwischen Ereignis und Ergebnis
- **Fallermittlung:** Auffinden von Personen mit bestimmten Erkrankungen durch Screening, Suchen in Krankenakten oder Befragung von Personen
- **Konfidenzintervall:** Ein Maß dazu, wie genau die Ergebnisse sind. Ein enges Konfidenzintervall verspricht Präzision, ein weites Ungenauigkeit.
- **Ereignis:** Der Faktor, von dem wir glauben, dass er das Krankheitsrisiko verändert.
- **Odds Ratio:** Das Verhältnis zwischen der Erkrankungswahrscheinlichkeit in einer Gruppe im Vergleich mit einer anderen. Odds werden genutzt, um Risiken abzuschätzen.
- **Ergebnis:** Die Krankheit oder der Umstand, der uns interessiert.
- **Rate:** Das Risiko, eine bestimmte Erkrankung innerhalb einer Zeitspanne zu erleiden.
- **Rate Ratio:** Das Verhältnis zwischen der Erkrankungsrate in einer Gruppe im Vergleich mit einer anderen
- **Wiedererkennungsfehler (Recall bias):** Ungenauigkeiten, die entstehen, wenn Personen (Fälle gegenüber Kontrollen) vergangene Ereignisse mit einem unterschiedlichen Ausmaß an Genauigkeit erinnern.
- **Residuelle Störfaktoren:** Störungen in schon adjustierten Ergebnissen, entweder weil wir einen bestimmten Faktor nicht kontrolliert haben oder weil ein Störfaktor ungenau gemessen wurde, sodass die Adjustierung den gesamten Effekt dieses Störfaktors nicht weggenommen hat.
- **Risikofaktor:** Ein Umstand, der mit einer Erkrankungswahrscheinlichkeit assoziiert ist, nicht notwendigerweise kausal.

Literatur

1. Okasha M. Epidemiological research. Student BMJ 2001; 9:277–8 (August)
2. Davey Smith G, Frankel S, Yarnell J. Sex and death: are they related? Findings from the Caerphilly cohort study. BMJ 1997; 315:1641–4
3. Sterne J, Davey Smith G. Sifting the evidence – what's wrong with significance tests? BMJ 2001;322:226–31
4. Hill AB. The environment and disease: association or causation? J R Soc Med 1965; 58:295–300

Weiterführende Literatur

Kirkwood BR. Essentials of medical statistics. Oxford: Blackwell Science, 1988
Last JM. A dictionary of epidemiology. Oxford: Oxford University Press, 1988