



Klinische Studien verstehen

Spurensuche

Bei der Fahndung nach der optimalen Behandlung geht der Arzt vielen Spuren nach. Wir verraten, auf was man beim Lesen klinischer Studien achten sollte und warum evidenzbasierte Medizin gut für die Patienten ist.

Jeder Patient hat das Recht, nach dem derzeit aktuellen medizinischen Wissensstand behandelt zu werden. Doch für jeden Einzelnen im Gesundheitswesen ist es kaum möglich, den Überblick über den aktuellen Wissensstand zu behalten. Ständig werden neue Geräte und Arzneimittel erprobt und die Resultate sind zum Teil widersprüchlich. Deshalb ist es wichtig, dass patientenorientierte Entscheidungen möglichst immer auf der Grundlage von nachgewiesener Wirksamkeit getroffen werden. Darauf basieren auch zum Beispiel die DMP.

Evidenzbasierte Medizin (EbM) ist die Richtung in der Medizin, die solche Wirksamkeitsnachweise mit Hilfe von statis-

tischen Verfahren erbringt. Der Begriff kommt vom englischen „evidence-based medicine“ und bedeutet soviel wie „auf wissenschaftlichen Belegen beruhende Behandlungsentscheidungen“. Grundlage sind klinische Studien, die einen genau beschriebenen Sachverhalt erhärten oder widerlegen. Die EbM beschäftigt sich nicht mit der Durchführung von solchen Studien, sondern allein mit der systematischen Nutzung ihrer Ergebnisse.

Diese Idee ist übrigens nicht neu. Schon 1753 wurden die Ergebnisse eines Versuchs veröffentlicht, Skorbut mit Orangen und Zitronen zu behandeln – was gelang, da Skorbut eine Vitamin-C-Mangelkrankung ist. Und der berühmte Ignaz Semmelweis entdeckte 1848 mit

einer „systematischen klinischen Beobachtung“ die Ursache des Kindbettfiebers und empfahl Desinfektion zur Vorbeugung. Evidenzbasierte Medizin soll dazu beitragen, die Anwendung unwirksamer oder gar schädlicher Verfahren zu vermeiden. Letztlich entscheidet der Arzt aber individuell, wie eine wissenschaftliche Erkenntnis sinnvoll für Patienten eingesetzt werden kann.

Was ist der Nutzen?

Um den Nutzen einer therapeutischen Maßnahme zu erforschen, beobachtet man wie sie wirkt. Dabei unterscheidet man experimentelle und beobachtende Studien. Beobachtende Studien können prospektiv (vorausschauend) oder retrospektiv (zurückblickend) sein, sie beobachten und protokollieren Zusammenhänge. Dadurch wurde zum Beispiel schon vor vielen Jahrzehnten der Zusammenhang von Rauchen und Lungenkrebs erkannt. Man erfasste bei einer Patientengruppe, die an Lungenkrebs erkrankt war, den Anteil der Patienten, die zuvor über viele Jahre geraucht hatten. Anschließend wurde in einer Vergleichsgruppe bei Menschen ohne Lungenkrebskrankung geschaut, wie hoch der Raucheranteil dort war. Der deutliche Unterschied beider Anteile ergab einen klaren Hinweis auf Rauchen als Hauptursache des Lungenkrebses.

Beobachtende Studien sind relativ einfach durchzuführen, haben aber den Nachteil, dass man auf viele Faktoren, die das Studienergebnis verfälschen können, keinen Einfluss hat. Bei experimentellen Studien, die grundsätzlich prospektiv erfolgen, versucht man dagegen, störende Einflüsse von vornherein auszuschließen. Um zu vermeiden, dass sich die Patienten in wichtigen Parametern unterscheiden (z.B. Alter, Geschlecht, Begleiterkrankungen), ordnet man die Studienteilnehmer nach dem Zufallsprinzip jeweils einer der beiden Gruppen zu (Randomisierung). Soll beispielsweise ein neues Medikament geprüft werden, besteht der einzige Unterschied im Idealfall dann darin, dass eine Gruppe das neue Medikament erhält (Interventions-

gruppe) und die andere nur ein Scheinmedikament oder das bislang bei der Erkrankung übliche Standardmedikament (Kontrollgruppe). In ähnlicher Weise wird auch die Wirksamkeit nichtmedikamentöser Diagnose- und Therapie-maßnahmen überprüft (z.B. Akupunktur versus Scheinakupunktur).

Um eine Verzerrung der Ergebnisse durch bestimmte Erwartungen der Forscher oder Patienten auszuschließen, wird zudem oft geheim gehalten, wer das neue Medikament und wer das Placebo- bzw. Standardpräparat bekommt. Bei einfachblinden Studien wissen nur die Patienten nicht, was sie erhalten, bei doppelblinden Studien wissen es auch die beteiligten Ärzte nicht. Die randomisierte, kontrollierte und doppelblinde Studie gilt heute als EbM-Standard.

Neben der Qualität einer Studie zählen natürlich auch ihre Ergebnisse. Ob der

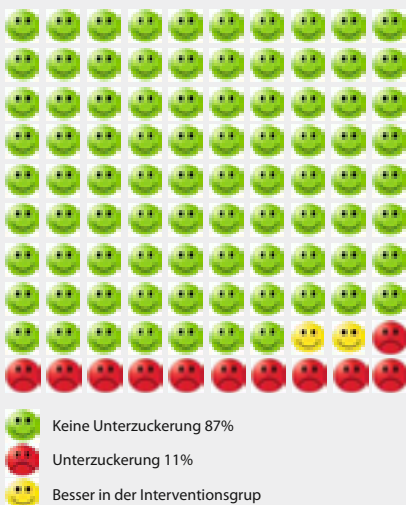
Unterschied zwischen einem Wirkstoff und Placebo nur auf einem Zufall beruht oder tatsächlich statistisch signifikant ist, das beschreibt der p-Wert. Je kleiner der p-Wert ist, desto geringer ist die Irrtumswahrscheinlichkeit. In der Regel spricht man bei einem p-Wert unter 0,05 von statistisch signifikanten Unterschieden.

Grenzen der Evidenz

Da sich statistische Ergebnisse sehr unterschiedlich interpretieren lassen (mehr darüber im Kasten), ist die Bewertung durch den Einzelnen oft schwierig – auch für den Arzt. In Deutschland übernimmt das deshalb der Gemeinsame Bundesausschuss bzw. das IQWiG: Neue Therapieverfahren und Medikamente werden nach EbM-Grundsätzen geprüft, und nur wenn die Belege für ihren Nutzen überzeugend sind, können sie in den Leistungskatalog der gesetzlichen Kranken-

versicherungen aufgenommen werden. Doch die „Evidenz“ hat auch Grenzen. Häufig liegen z.B. nur Studien zu willkürlich gewählten „Zwischen“-Ergebnissen vor (z.B. Blutzuckerspiegel oder HbA1c-Wert bei Diabetikern), die nur begrenzt Aussagen zu den für die Patienten relevanten klinischen Ergebnissen (z.B. Sterblichkeit oder Herzinfarkttrate) erlauben. Und bei Kindern sind Studien erfahrungsgemäß schwierig durchzuführen, weil die kleinen Patienten viele Dinge nicht beurteilen können und ihr Körper sehr empfindlich auf Wirkstoffe reagieren kann. Andere Sachverhalte sind dagegen so selbstverständlich, dass es kaum Untersuchungen dazu gibt. So hat eine Studie von 1954 als erste und letzte die Entstehung von Karies durch Zucker untersucht. Das Ergebnis zeigte so eindeutig die Auswirkungen des Zuckers, dass alle überzeugt waren. **RM ■**

Studienmathematik für Interessierte



Durch Behandlung von 100 Patienten lassen sich zwei Unterzuckerungen vermeiden (NNT = 50).

Wenn ein neues Arzneimittel getestet wird, werden klare Aussagen erwartet. Zum Beispiel, ob das Risiko für eine Folgekrankheit oder unerwünschte Wirkungen der bisherigen Therapie dadurch reduziert wird oder nicht. Es gibt aber verschiedene Möglich-

keiten der Interpretation: Das **relative Risiko** beschreibt das Verhältnis der Wahrscheinlichkeit für ein bestimmtes Ereignis zwischen Interventions- und Kontrollgruppe. Ein relatives Risiko von 1 heißt kein Unterschied zwischen den Gruppen, ein relatives Risiko unter 1 zeigt, dass die Behandlung ein Risiko reduzieren kann. Ist das Risiko in der Interventionsgruppe 1% und in der Kontrollgruppe 2%, so beträgt die relative Risikoreduktion (RRR) 50%. Das **absolute Risiko** (ARR) reduziert sich aber nur um 1%. Hochwertige Studien errechnen deshalb oft die Anzahl von Patienten, die behandelt werden müssen, um ein erwünschtes Ereignis zu erzielen oder ein unerwünschtes zu verhindern. Diese Zahl wird als „**number needed to treat**“ (NNT) bezeichnet. Diese „Anzahl notwendiger Behandlungen“ gilt immer nur für diese eine Behandlung und ist ein gutes Maß für die Wirksamkeit. Die NNT errechnet sich aus der Formel $1/ARR * 100$.

Ein fiktives Rechenbeispiel: 13% der Diabetiker, die mit einem gängigen blutzucker-

senkenden Medikament behandelt wurden, erlitten eine unerwünschte Unterzuckerung. Bei Patienten, die mit einem neuen Medikament therapiert wurden, passierte das nur in 11% der Fälle. Die RRR beträgt dann 15,3%, die ARR nur 2% und die NNT 50 ($1/2*100$). Das bedeutet, dass 50 Diabetiker mit dem neuen Medikament behandelt werden müssten, um eine Unterzuckerung zu verhindern. Schauen Sie also genau hin. Zunächst überzeugt die relative Risikoreduktion: mehr als 15% weniger Unterzuckerungen! Die absolute Risikoreduktion ist mit 2% schon weniger beeindruckend und die NNT drückt aus, mit welchem Aufwand man diese Senkung des Risikos erzielen kann. Das Buch „Einmaleins der Skepsis“ von Gerd Gigerenzer stellt viele solche Beispiele vor, auch aus ganz anderen Anwendungen: Wenn etwa durch zweijähriges Mammografie-Screening das Leben einer von 1.000 teilnehmenden Frauen gerettet wird, so ist die NNT gleich 1.000. Die übrigen 999 Frauen haben keinen Nutzen im Sinne einer Sterblichkeitsverminderung.