



Evidenzbasierte Medizin

Gute Studien, schlechte Studien

Warum werden viele Behandlungen von der GKV erstattet, manche anderen aber nicht? In der Regel hat das damit zu tun, ob es Nachweise für die Wirksamkeit der Behandlung gibt oder nicht. Wir erklären die Grundregeln der evidenzbasierten Medizin und den Unterschied zwischen guten und schlechten Studien.

Wir dürfen in Deutschland zurecht stolz sein auf unser Gesundheitssystem: Jeder Patient hat das Recht, nach dem derzeit aktuellen medizinischen Wissensstand behandelt zu werden. Doch was ist der aktuelle medizinische Stand? Ständig werden neue Geräte und Arzneimittel erprobt, und die Resultate sind zum Teil sehr widersprüchlich. Entscheidungen im Sinne der Patienten sollten deshalb immer auf der Grundlage von nachgewiesener Wirksamkeit getroffen werden. Das ist die evidenzbasierte Medizin (EbM).

Sie basiert auf Studien, in denen es mindestens zwei Gruppen gibt, die miteinander verglichen werden. Den Nutzen eines neuen Medikaments kann man so prüfen, indem man freiwillige Teilnehmerinnen und Teilnehmer nach dem Zufallsprinzip auf mindestens zwei Gruppen aufteilt („randomisiert“): Die eine Grup-

pe erhält das neue Medikament, die andere das bisher eingesetzte Medikament (Goldstandard) oder ein Placebo. Anschließend werden die Behandlungsergebnisse verglichen.

Dabei sollte der individuelle Einfluss der Ärzte sowie auch der Patienten so klein wie möglich gehalten werden. In hochwertigen Studien wird die Zuordnung deshalb „doppelt verblindet“: Dann wissen weder die Teilnehmenden noch die behandelnden Ärztinnen und Ärzte, wer zu welcher Gruppe gehört.

Viele verschiedene Studienarten

Um verlässliche Ergebnisse liefern zu können, muss eine Studie vor Beginn sorgfältig geplant werden. Dabei ist es besonders wichtig, die richtige Studienart auszuwählen. Der wichtigste Studientyp sind **randomisierte kontrollierte Studien** (engl. „randomized controlled

trial, RCT). Sie werden z. B. durchgeführt, wenn die Wirkung eines neuen Medikamentes geprüft werden soll. Im Vorfeld wird bestimmt, was genau untersucht werden soll und an welcher Personengruppe. Das Akronym PICO steht dabei für die vier wichtigen Punkte, hier dargestellt an einem Beispiel:

- > **P** (Patient oder Problem: wer wird behandelt?), z. B. Patienten mit Diabetes mellitus Typ I
- > **I** (Intervention: womit wird behandelt?), z. B. Insulinanaloge
- > **C** (Comparison oder Control: was ist der Vergleich?), z. B. Insulin
- > **O** (Outcome: was soll erreicht werden?), z. B. Lebensqualität

Damit Aussagen über den Erfolg der Behandlung möglich sind, wird zudem geklärt, über welchen Zeitraum die Studie laufen muss und wie viele Teilnehmer benötigt werden.

Ein anderer Studientyp sind **Kohortenstudien**. Sie beobachten zwei (oder mehr) Gruppen, die verschiedenen Einflüssen ausgesetzt sind, etwa Raucher und Nichtraucher. Dann kann untersucht werden, wie sich der Gesundheitszu-

stand in beiden Gruppen über die Jahre entwickelt, ob und woran die Teilnehmer erkranken und wie viele von ihnen sterben. An Kohortenstudien nehmen oft Menschen teil, die zu Studienbeginn gesund sind. Dann spricht man von einer prospektiven Studie. Die zu messenden Ereignisse und andere mögliche Einflussgrößen werden im Vorfeld festgelegt. In einer retrospektiven Studie ist die Erkrankung zu Beginn der Studie schon eingetreten, und es wird rückblickend nach Risikofaktoren gesucht.

In **Fall-Kontroll-Studien**, einem dritten Studientyp, werden kranke Menschen mit anderen Personen verglichen, die die Erkrankung nicht haben, ihnen aber ansonsten möglichst ähnlich sind (z. B. gleiches Geschlecht und Alter). Kohortenstudien und Fall-Kontroll-Studien werden zusammen auch als Beobachtungsstudien bezeichnet. Der letzte Typ, die **Querschnittstudie**, ähnelt mehr einer Umfrage. Eine Zufallsstichprobe wird befragt oder untersucht, um bestimmte Fragen zu beantworten. „Wie viele Menschen nehmen an einer Krebsfrüherkennungsuntersuchung teil?“, könnte eine solche Fragestellung lauten. Diese Studien sind schneller und günstiger, erlauben jedoch keine Aussagen darüber, was eine Erkrankung verursacht oder wie sie am besten behandelt wird. Verlässliche Aussagen über Ursache und Wirkung sind nur aufgrund von Ergebnissen randomisierter kontrollierter Studien möglich.

Wie werden Studien bewertet?

Um zu bewerten, ob eine Studie verlässliche Daten liefert, muss zunächst einmal geprüft werden, weshalb sie überhaupt gemacht worden ist und welche Frage-

stellung sie untersucht hat. Waren die Kriterien, die in der Studie als wichtige Zielgröße dienen, richtig gewählt? Gerne werden Laborwerte als Zielparameter genommen, etwa Blutfette oder Blutzucker. Manchmal verändern sich solche Laborwerte positiv, die Patienten haben aber trotzdem keinen Vorteil von der Behandlung. Deshalb unterscheidet man zwischen **patientenrelevanten Parametern**, die für erkrankte Menschen entscheidend sind (ihr Leben verlängert, ihnen Klinikaufenthalte erspart, ihre Beschwerden verringert, Komplikationen vorbeugt oder ihren Alltag und den Umgang mit der Erkrankung erleichtert) und sogenannten **Surrogatparametern** (vom lateinischen *surrogatum* = der Ersatz).

Surrogate sind deutlich einfacher zu erheben: Man sieht in einer Studie schnell, ob ein Medikament beispielsweise den Blutdruck senkt. Es kann jedoch Jahre dauern, bis sich sagen lässt, ob dadurch auch Erkrankungen wie Herzinfarkte verhindert werden. Das hat statistische Gründe: Herzinfarkte sind relativ selten, sodass man eine große Zahl von Menschen beobachten muss, um eindeutige Unterschiede zwischen verschiedenen Behandlungsgruppen sehen zu können. In einer Studie zum Blutdruck sind nur wenige Teilnehmer nötig, um eine Wirkung festzustellen, denn man kann bei jedem von ihnen messen, wie sich der Blutdruck verändert. Es ist also Vorsicht geboten, wenn Labor- und Körpermesswerte in Studien als Surrogate verwendet werden, um den Nutzen einer Therapie für Patientinnen und Patienten zu bestimmen.

Metaanalysen

Fasst man die Ergebnisse aller gefundenen Studien zu einem Thema zu einem Gesamtergebnis zusammen, nennt man das eine **systematische Übersicht** oder **Meta-Analyse**. Das Gesamtergebnis hat oft eine deutlich höhere Aussagekraft als die Ergebnisse der Einzelstudien. Eine solche Auswertung ist aber nur sinnvoll, wenn die Ergebnisse der einzelnen Stu-



Die Pyramide beschreibt die Wertigkeit der einzelnen Studientypen von oben nach unten. Quelle: www.eupati.eu

dienergebnisse relativ ähnlich (homogen) sind. Wenn sich die Ergebnisse stark unterscheiden, ist das ein Hinweis, dass die Unterschiede zwischen den Studien genauer untersucht werden sollten, damit man nicht Äpfel mit Birnen vergleicht.

Die Cochrane Library

Eine wichtige Quelle für systematische Übersichten ist zum Beispiel die Datenbank „Cochrane Library“. Sie wird herausgegeben von einem internationalen Wissenschafts-Netzwerk, das sich auf solche Übersichten spezialisiert hat. Auch der Gemeinsame Bundesausschuss (G-BA) greift gerne auf solche systematischen Übersichten zurück, wenn er bestimmte Präventions-, Diagnose- oder Therapieverfahren beurteilen soll. Experten aus den medizinischen Fachgesellschaften, von niedergelassenen Ärzten und Krankenhäusern sowie Patientenverbänden prüfen dort die Evidenz. Bei fehlendem Nachweis der Wirksamkeit oder wenn sich keine patientenrelevanten Endpunkte ergeben, wird das getestete Verfahren mit hoher Wahrscheinlichkeit keinen Einzug in den Leistungskatalog der GKV erhalten. Das schützt Patienten vor nutzlosen und möglicherweise gefährlichen Behandlungen und spart dem Gesundheitssystem Geld, das an anderer Stelle vernünftiger eingesetzt werden kann.

Webtipps
 Informationsportal des IQWiG mit umfassenden Beiträgen zu Studien und ihrer Auswertung:
www.gesundheitsinformation.de
 > Rubrik: Evidenzbasierte Medizin
 Informationen der AOK für Praxisteam zum Thema:
www.aok-gesundheitspartner.de,
 Webcode W128324