

Bestimmung der Angemessenheit bei medizinischen Leistungen

Indikationsqualität: leicht gesagt, schwer zu fassen

Johann Steurer

Prof. em. Dr. med., Universität Zürich, Mitglied Ausschuss Qualitätsboard, Universitätsspital Zürich

In den letzten Jahren lag der Fokus der Qualitätskontrolle und Qualitätssicherung für medizinische Leistungen vorwiegend auf der Outcome-Qualität. Diese gilt erfreulicherweise insgesamt als sehr gut. Allerdings ist bis jetzt unklar, ob die getätigten Messungen tatsächlich zur Qualitätsverbesserung beitragen. Deshalb werden vermehrt Forderungen nach der Indikationsqualität laut. Doch deren Erfassung ist nicht ganz einfach.

Die Gesellschaft wünscht sich eine qualitativ hochstehende medizinische Betreuung. Die Patientinnen und Patienten erwarten von den Medizinerinnen und Mediziner, dass sie jeder Person die für ihre Situation angemessene Untersuchung und Behandlung empfehlen und diese korrekt durchführen. In den letzten Jahren lag der Fokus der Qualitätskontrolle und Qualitätssicherung vorwiegend auf der Outcome-Qualität (Ergebnisqualität).

Viele auch öffentlich zugängliche Daten zu diesen mit grossem Aufwand und hohen Kosten erfassten Outcome-Messungen sind vorhanden, und die Qualität gilt generell als sehr gut. Es gibt aber wenig Hinweise, dass sich die Outcome-Qualität infolge der vielen Messungen und qualitätssichernden Aktivitäten in den vergangenen Jahren verbessert hat.

Neuer Fokus: Indikationsqualität

Die politischen Behörden verlangen in Zukunft nicht nur Angaben zur Outcome-Qualität, sondern auch Angaben zur Indikationsqualität. Gemeint ist damit die Angemessenheit einer Untersuchung oder einer Therapie. Der Anteil angemessener Untersuchungen und

In der Medizin gibt es Situationen, in denen klar ist, was angemessen ist, doch bei der Mehrzahl der Leistungen ist es nicht eindeutig.

Behandlungen an der Gesamtheit der erbrachten medizinischen Leistungen liefert einen Hinweis auf die Indikationsqualität einer Ärztin resp. eines Arztes, einer Klinik oder eines Spitals.

Je nach Bericht sind 10 bis 20% der in der Schweiz erbrachten diagnostischen und therapeutischen Leis-

tungen nutzlos oder nur von geringem Nutzen [1]. Diese unnötigen Leistungen verursachen jährlich Ausgaben in Milliardenhöhe. Teilweise sind sie für die Patientinnen und Patienten unangenehm und potenziell auch schädlich. Dies erklärt die Forderung der Behörden, dass medizinische Leistungen angemessen sein sollten, und es begründet ihr Bedürfnis, die Indikationsqualität zu erfassen und die Ergebnisse zu ver-

10 bis 20% der in der Schweiz erbrachten diagnostischen und therapeutischen Leistungen haben keinen oder nur geringen Nutzen.

öffentlichen. Voraussetzungen für die Bewertung der Indikationsqualität sind:

1. Festlegen von Kriterien, um die Verhältnismässigkeit zu beurteilen;
2. Sammeln von Daten, um zu prüfen, ob die Kriterien erfüllt sind oder nicht.

In der Medizin gibt es Situationen, in denen klar ist, was angemessen ist, und es gibt Situationen – das trifft auf die Mehrzahl der Leistungen zu –, in denen das nicht so klar ist. Für einen Teil der medizinischen Leistungen ist es einfach, Kriterien zur Beurteilung der Indikationsqualität zu erstellen und zu erfassen, für andere Leistungen ist das aber nicht der Fall.

Generell ist eine Leistung angemessen, wenn entweder bei einer Person die diagnostischen Untersuchungen eine vermutete Krankheit mit ausreichend hoher Wahrscheinlichkeit ausschliessen bzw. nachweisen können. Oder sie ist angemessen, wenn eine Therapie den Verlauf einer diagnostizierten Krankheit (den Outcome) verbessert.



Bei vielen medizinischen Leistungen ist es schwierig, einheitliche Kriterien zur Messung der Indikationsqualität festzulegen.

Bekannte Grenzfälle

Es gibt Untersuchungen, zu deren Angemessenheit ein Konsens zwischen den Expertinnen und Experten besteht. Beispiele dafür sind: EKG und Troponin-Bestimmung bei einer Person mit Verdacht auf einen Myokardinfarkt, ein CT der Lunge bei dringendem Verdacht auf eine Lungenembolie oder ein Schädel-CT bei einer Patientin bzw. einem Patienten mit neu aufgetretenen starken Kopfschmerzen und getrübtetem Bewusstseinszustand. Diese Untersuchungen dienen dem Nachweis oder Ausschluss gravierender Krankheiten. Wenn diese Krankheiten nicht rechtzeitig diagnostiziert und behandelt werden, sind die Folgen für die Betroffenen schwerwiegend.

Eindeutig nicht indiziert

Es gibt auch diagnostische Tests, für die gemäss der «smarter medicine» (Choosing Wisely Switzerland) Konsens besteht, dass sie *nicht indiziert* sind.

Beispiele dafür sind:

- Bestimmung von Vitamin D im Blut, wenn keine Risiken für einen Mangel (z.B. geringe Sonnenexposition, entzündliche Darmerkrankungen) vorliegen.
- Bestimmung einer Palette von Autoantikörpern (Antikörper gegen dsDNA, RNP, SSA SSB, Scl70 und andere), wenn der ANA-Test (antinukleäre Antikörper) negativ ist.
- Bestimmung von Antikörpern gegen Borrelien bei Personen ohne spezifische Symptome und Befunde, die auf eine Infektion damit hindeuten.

Ein weiteres Beispiel: Bei Patientinnen und Patienten mit unspezifischen Rückenschmerzen und ohne neurologische Ausfallzeichen ist in den ersten sechs Wochen nach Beginn der Schmerzen ein bildgebendes Verfahren (konventionelles Röntgenbild, CT, MRT) nicht indiziert [2].

Diese Untersuchungen sind, abgesehen von wenigen Ausnahmen, nicht angezeigt, da ihre Resultate keinen oder nur einen marginal relevanten Wissenszuwachs generieren. In der Praxis ist es nicht immer einfach, diese Untersuchungen so zu klassifizieren.

Beim Beispiel bei der Antikörperbestimmung gegen Borrelien heisst es in den Empfehlungen, dass die Untersuchung bei Fehlen spezifischer Symptome nicht indiziert ist. Die Frage ist, was spezifische Symptome sind. Die Borrelieninfektion manifestiert sich mit unterschiedlichen Symptomen, und welche davon spezifisch sind und welche nicht, ist nicht einfach festzulegen.

Weitere Knackpunkte

Weiter gibt es diagnostische Tests, bei denen die Beurteilung ihrer Angemessenheit schwieriger ist. Ist bei einem Achtzigjährigen mit Schnupfen, Husten und einer seit drei Tagen leicht erhöhten Körpertemperatur ein Thoraxröntgenbild indiziert, um eine Pneumonie als Ursache der Beschwerden auszuschliessen? Ist ein Röntgenbild indiziert, wenn der CRP-Wert über 50 mg/l, oder erst, wenn er über 80 mg/l liegt? Ist bei einer Patientin mit einer stabilen Angina pectoris eine Koronar-Angiographie indiziert? Gemäss

den Empfehlungen der European Society of Cardiology [3] ist eine Koronar-Angiographie ohne vorheriges Belastungs-EKG oder Stress-Echo indiziert, wenn die klinische Wahrscheinlichkeit aufgrund der Anamnese hoch (wie hoch?) ist. Bei nicht so hoher klinischer Wahrscheinlichkeit wird vor einer Koro-

Es ist bekannt, dass es selbst für erfahrene Medizinerinnen und Mediziner schwierig ist, Krankheitswahrscheinlichkeiten zu schätzen.

nar-Angiographie ein Belastungstest empfohlen. Abhängig von dessen Ergebnis gilt eine Koronar-Angiographie als angemessen oder eben als nicht angemessen.

Wenn man Kriterien für die Angemessenheit definiert, stellen sich bei dieser vereinfachten Schilderung zwei Herausforderungen. Zum einen stellt sich die Frage, welche klinische Wahrscheinlichkeit als «hoch» gelten soll, zum anderen stellt die Verlässlichkeit der Schätzung klinischer Wahrscheinlichkeiten Probleme.

Es ist bekannt, dass es selbst für erfahrene Medizinerinnen und Mediziner schwierig ist, Krankheitswahrscheinlichkeiten zu schätzen. Bei verschiedenen Studien lagen die Schätzungen der Fachpersonen für ein und denselben Fall nicht um ein paar Prozentpunkte auseinander, sondern die einen schätzten eine Wahrscheinlichkeit von 10% oder tiefer, die anderen von 90% oder höher [4].

Unsicherheiten herrschen auch bei der Durchführung einer Ultraschalluntersuchung des Abdomens. Es besteht Konsens, dass sie bei Verdacht auf eine akute Appendizitis oder bei Verdacht auf einen Stein im Gallengang oder in den ableitenden Harnwegen indiziert ist, ob sie aber auch bei «leichten» Bauchschmerzen indiziert ist, lässt sich nicht wirklich beantworten.

Und bei der Therapie?

Das Ziel einer Therapie ist es, den Verlauf einer Krankheit – erfasst mit einem oder mehreren Outcome-Parametern – positiv zu beeinflussen. Damit soll die Heilung, die Linderung von Symptomen und Schmerzen, die Wiederherstellung körperlicher Funktionen oder die Verhinderung einer weiteren Verschlechterung körperlicher Funktionsfähigkeiten erreicht werden.

Für viele Krankheiten wurden und werden *Guidelines* (Empfehlungen) verfasst, die den Fachpersonen helfen sollen, den Patientinnen und Patienten die bestmögliche Behandlung zu empfehlen. Bei Therapieempfehlungen geht es um das Abwägen des potenziellen

Nutzens einer Therapie gegenüber ihrem potenziellen Schaden (Nebenwirkungen) – und das immer im Vergleich mit anderen Therapieoptionen oder dem Verzicht auf eine Therapie.

Unbestrittene Fälle

Es gibt einige Krankheiten, für die es nur eine Therapieoption gibt. So ist bei fast allen Personen mit einer Fraktur des Schenkelhalses eine chirurgische Therapie indiziert. Das Gleiche gilt für die Therapie, ob chirurgisch oder endoskopisch, von Personen mit einem Stein, der den Gallengang obstruiert. Es besteht auch Konsens, dass Patientinnen und Patienten mit einer Hypertonie oder Diabetes mit den entsprechenden Medikamenten oder Massnahmen behandelt werden sollten.

Ein Unterlassen der Therapie, wenn sie auch nicht bei allen Betroffenen zum erwarteten Erfolg führt, hat für die Betroffenen in einen Fall sofort, im anderen Fall in

Es gibt diagnostische Tests, für die gemäss «smarter medicine» Konsens besteht, dass sie nicht indiziert sind.

ein paar Jahren oder Jahrzehnten gravierende Folgen. Es sind dies Situationen, in denen der potenzielle Nutzen der Therapie im Vergleich mit keiner Therapie für *fast alle* Betroffenen grösser ist als der mögliche Schaden. In diesen Situationen ist es einfach, Kriterien zu definieren, um die Indikationsqualität zu beurteilen.

Beispiele für eine schwierige Bewertung

Es gibt aber Krankheiten, bei denen der potenzielle Nutzen der Therapie, verglichen mit keiner Therapie, nur für einen *unterschiedlich grossen Teil* der Patientinnen und Patienten grösser ist als der mögliche Schaden.

Erweiterung Spinalkanal bei Lumbalstenose

Ein Beispiel dafür: Personen mit einer Lumbalstenose profitieren von einer chirurgischen Erweiterung des engen Spinalkanals. Das ist das Ergebnis klinischer Studien, in denen ein Teil der Patientinnen und Patienten operiert wurde und ein anderer Teil konservativ, mit Physiotherapie oder gar nicht behandelt wurde [5]. Die Ergebnisse klinischer Studien repräsentieren einen Mittelwert der Wirksamkeit im untersuchten Patientenkollektiv, erlauben aber keine Aussagen zur Wirksamkeit für eine Einzelperson.

Klinische Studien zeigen, dass sich bei zwei Drittel der Behandelten die Beschwerden nach der Operation deutlich bessern, bei einem Drittel ist keine Besserung zu beobachten, und bei einigen nehmen die Beschwerden nach der Operation sogar zu [6]. Vor der Operation ist es nicht möglich, eine verlässliche Aussage darüber zu treffen, ob die Person von der Operation profitieren wird oder nicht.

Die Frage ist, ob eine Operation bei allen Patientinnen und Patienten (Personen mit zusätzlichen Krankheiten und Kontraindikationen ausgenommen) mit moderaten Beschwerden angemessen ist oder nicht; bei starken Beschwerden besteht Konsens für eine Operation.

Unbeantwortet ist auch die Frage, ob bei der Operation ein interspinaler Abstandhalter («Spacer») implantiert werden soll oder nicht. Zu diesen Fragen besteht kein Konsens, und die Beurteilung der Indikationsqualität ist somit schwierig bis unmöglich.

Ruptur des Kreuzbandes

Ein weiteres Beispiel, an dem die Schwierigkeit, Angemessenheitskriterien zu definieren, gezeigt werden kann, ist die Behandlung von Patientinnen und Patienten mit einer Ruptur des vorderen Kreuzbandes. Die Autoren einer randomisierten Studie kommen zum Schluss, dass eine operative Therapie bezüglich des funktionellen Outcomes nach zwei und fünf Jahren keinen Vorteil gegenüber einer konservativen Therapie zeigt und 50% der untersuchten Personen auch ohne Operation (also nur mit Physiotherapie) ein «gutes» Ergebnis zeigen [7].

Eine andere Studie [8] kommt zum Schluss, dass eine konservative Behandlung bei vielen Personen nach einem Jahr ein «gutes» Ergebnis zeigt. In der operierten Gruppe ist der Anteil der Patientinnen und Patienten mit einem guten Ergebnis aber grösser. Die Hälfte der Betroffenen, die primär konservativ behandelt worden waren, wurde im Verlauf von zwei Jahren doch operiert. Wie bei der Lumbalstenose kann auch bei der Ruptur des vorderen Kreuzbandes vor der Operation keine verlässliche Aussage getroffen werden, wer von der Operation profitiert und wer nicht. Ausserdem sind bei der Festlegung von Angemessenheitskriterien auch die Kosten der Behandlung in Betracht zu ziehen. So belaufen sich laut den Berech-

Das Festlegen von Angemessenheitskriterien für die Einzelperson – und um sie geht es in der klinischen Medizin – ist herausfordernd.

nungen des *Medical Board* die Kosten bei einer operativen Behandlung der Kreuzbandruptur auf rund CHF 670 000 für ein zusätzliches Lebensjahr bei optimaler Lebensqualität (QALY) [9].

Die beiden Beispiele zeigen, dass das Festlegen von Angemessenheitskriterien für die individuelle Person – und um sie geht es in der klinischen Medizin – eine herausfordernde Aufgabe ist. Es ist nicht damit getan, aufgrund einer oder mehrerer Studienergebnisse (Metaanalysen) zu entscheiden, ob für eine Patientin bzw. einen Patienten eine Therapie angemessen ist oder nicht.

Was zu tun ist

Wenn Politik und Behörden die Erfassung der Angemessenheit medizinischer Leistungen, die Indikationsqualität, fordern, darf das nicht alle, sondern höchstens eine Auswahl medizinischer Leistungen betreffen. Vernünftig wäre es, Leistungen auszuwählen, für die es Hinweise gibt, dass ein relevanter Teil davon nicht angemessen ist und sich die Fachpersonen mehrheitlich auf Angemessenheitskriterien einigen können.

Für diese Leistungen sind situationsbezogene (nicht «one size fits all») und überprüfbare Kriterien zu formulieren. Dies ist in erster Linie die Aufgabe medizinischer Fachgesellschaften.

Die Daten, die für die Beurteilung der Angemessenheit medizinischer Leistungen notwendig sind, müssen strukturiert erfasst und standardisiert gespeichert werden. Mit den meisten heute verfügbaren Klinikinformationssystemen bzw. elektronischen Patientendossiers ist das nicht möglich.

Der Aufwand für die Erarbeitung der Kriterien und die Erfassung der dazu notwendigen Daten ist beträcht-



Bei der Bestimmung von Angemessenheitskriterien sollten auch die Kosten für die Behandlung miteinbezogen werden.

lich. Die dabei anfallenden Kosten können nicht auf die Spitäler abgewälzt werden. Die Steuerzahlerinnen und Steuerzahler oder die Krankenkassen (also die Versicherten) müssen diese Mehrkosten tragen. Der potenzielle Nutzen der Erfassung der Indikationsqualität ist gegen die entstehenden Kosten abzuwägen. Und nicht zuletzt: Eine hohe Indikationsqualität muss sich für die Leistungserbringer lohnen. Wenn nur noch angemessene Leistungen erbracht würden, also die beträchtliche Menge an nutzlosen Leistungen massiv reduziert würde, würde die Zahl medizinischer Leistungen ebenfalls stark sinken – und das lohnt sich für die Leistungserbringer nicht.

Bildnachweis

patricia serna / Unsplash
Megaflopp | Dreamstime.com

Literatur

1 Bericht der Expertengruppe. Kostendämpfungsmaßnahmen zur Entlastung der obligatorischen Krankenpflegeversicherung. 2017. Available from: <https://www.bag.admin.ch/bag/de/home/versicherungen/krankenversicherung/kostendaempfung-kv.html>

2 smarter medicine – Choosing Wisely Switzerland. Rheumatologie Available from: <https://www.smartermedicine.ch/de/top-5-listen/rheumatologie.html>

3 Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, et al. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes. Eur Heart J. 2020;41(3):407–77.

4 Steurer J, Held U, Miettinen OS. Diagnostic probability function for acute coronary heart disease garnered from experts' tacit knowledge. Journal of clinical epidemiology. 2013;66(11):1289–95.

5 Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B, et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. N Engl J Med. 2008;358(8):794–810.

6 Held U, Burgstaller JM, Wertli MM, Pichierri G, Winklhofer S, Brunner F, et al. Prognostic function to estimate the probability of meaningful clinical improvement after surgery – Results of a prospective multicenter observational cohort study on patients with lumbar spinal stenosis. PLoS One. 2018;13(11):e0207126.

7 Frobell RB, Roos HP, Roos EM, Roemer FW, Ranstam J, Lohmander LS. Treatment for acute anterior cruciate ligament tear: five year outcome of randomised trial. BMJ. 2013;346:f232.

8 Reijman M, Eggerding V, van Es E, van Arkel E, van den Brand I, van Linge J, et al. Early surgical reconstruction versus rehabilitation with elective delayed reconstruction for patients with anterior cruciate ligament rupture: COMPARE randomised controlled trial. BMJ. 2021;372:n375.

9 Gesundheitsdirektion des Kantons Zürich. Ruptur des vorderen Kreuzbandes: operative oder konservative Behandlung. Bericht vom 30. Juni 2009. https://www.swissmedicalboard.ch/fileadmin/public/news/2011_und_aelter/bericht__smb_kreuzbandriss_lang_2009.pdf

johann.steurer[at]jusz.ch

Das Wichtigste in Kürze

- Bisher wurde zur Qualitätsmessung und -sicherung von medizinischen Leistungen vor allem die Outcome-Qualität untersucht. Neu soll verstärkt die Indikationsqualität gemessen werden.
- Obwohl bei einigen diagnostischen Untersuchungen und Therapiemethoden bereits bekannt ist, dass sie häufig nicht angemessen sind, hängt die Bestimmung immer vom individuellen Einzelfall ab.
- Das macht die Festlegung von verbindlichen Angemessenheitskriterien schwierig. Sie sollten sich auf Leistungen beschränken, bei denen bereits Hinweise auf eine fehlende Angemessenheit bestehen.

L'essentiel en bref

- Jusqu'à présent, la mesure et l'assurance de la qualité des prestations médicales dépendaient essentiellement de la qualité des résultats. La qualité des indications devrait être davantage prise en compte.
- S'il est connu que certains examens diagnostiques et méthodes thérapeutiques ne sont souvent pas appropriés, le choix d'y recourir se décide toujours au cas par cas.
- Il est ainsi difficile de définir des critères fixes permettant de faciliter ce choix. Ces critères devraient se limiter aux prestations dont on sait qu'elles sont souvent inappropriées.