

Wahrscheinlichkeit und Unsicherheit – zwei Seiten derselben Medaille

Prof. em. Dr. med. Johann Steurer

In der Medizin wird kontinuierlich in Wahrscheinlichkeiten gedacht und gesprochen. Aufgrund des Alters (68 Jahre) und der vorliegenden Anamnese – plötzlich aufgetretene linksthorakale Schmerzen mit Ausstrahlung in den Hals und den linken Oberarm – wird bei Herrn Gabathuler ein Herzinfarkt als sehr wahrscheinlich angesehen. Demgegenüber erscheint eine Lungenentzündung bei Frau Inglin, die seit drei Tagen an Husten leidet und nur leicht erhöhte Temperaturen zeigt, als eher unwahrscheinlich. Ein unauffälliger D-Dimer-Wert lässt mit grosser Wahrscheinlichkeit eine Lungenembolie ausschließen. Ein erhöhter Troponinwert bei einem Patienten mit Brustschmerzen lässt einen Herzinfarkt als sehr wahrscheinlich erscheinen. Die regelmäßige Einnahme eines Statins nach einem Herzinfarkt reduziert das Risiko eines erneuten Infarkts in den kommenden sechs Jahren um etwa 6 Prozent (1).

In der medizinischen Praxis werden Wahrscheinlichkeiten typischerweise in qualitativen Kategorien wie zum Beispiel "unwahrscheinlich" oder "sehr wahrscheinlich" beschrieben, wobei konkrete Zahlenangaben eher selten sind.

Dem Begriff der Wahrscheinlichkeit steht der Begriff der Unsicherheit komplementär gegenüber. Die beiden Begriffe beschreiben zwei Seiten derselben Medaille: Wahrscheinlichkeit ist das Instrument mit dem Unsicherheit im klinischen Alltag geschätzt (praktisch nie berechnet) und kommuniziert wird. Sir W. Osler, ein renommierter Mediziner, der von 1849 bis 1919 lebte und als Wegbereiter der modernen Medizin gilt, hat den Aphorismus "Medicine is a science of uncertainty and an art of probability" geprägt.

Das Verständnis von Wahrscheinlichkeiten und ihre korrekte Interpretation spielen in der Medizin eine zentrale Rolle. Wenn ich an meine Ausbildung zurückdenke – die allerdings schon einige Jahre zurückliegt – erinnere ich mich daran, dass uns Wahrscheinlichkeiten vor allem anhand klassischer Beispiele vermittelt wurden: Wie wahrscheinlich ist es, dass beim Münzwurf Kopf oder Zahl erscheint? Wie oft fällt beim Würfeln eine Fünf? Oder: Mit welcher Wahrscheinlichkeit folgt auf eine geworfene Fünf eine Drei? Anspruchsvoller wurde es, wenn wir berechnen sollten,

wie wahrscheinlich beim gleichzeitigen Wurf zweier Würfel eine Drei und eine Vier auftreten. Das Konzept der «bedingten Wahrscheinlichkeiten» wurde so umständlich und langweilig erklärt, dass sich das Interesse der Studentinnen und Studenten an diesem Thema in Grenzen hielt (weitere Ausführungen zu bedingten Wahrscheinlichkeiten am Schluss des Artikels).

Bis heute ist mir jedoch unklar, welchen praktischen Nutzen solche Beispiele – mit Ausnahme der bedingten Wahrscheinlichkeiten – für die ärztliche Tätigkeit haben. In über zwanzig Jahren klinischer Praxis habe ich weder gewürfelt noch eine Münze geworfen, um eine Entscheidung zu treffen.

In der medizinischen Praxis stellen sich andere Fragen – etwa:

- Wie hoch ist die Wahrscheinlichkeit – qualitativ (z. B. eher unwahrscheinlich) oder quantitativ (z. B. 20 %) – oder anders gefragt, wie sicher oder unsicher bin ich, dass Herr Gabathuler an einer Lungenentzündung leidet?
- Wie wahrscheinlich ist es, dass Frau Kunz nach ihrer Rückenoperation wieder schmerzfrei gehen kann?

Wenn ich antworte, Herr Gabathuler habe mit einer Wahrscheinlichkeit von 20 % eine Lungenentzündung, und Frau Kunz werde mit 80 % Wahrscheinlichkeit nach der Operation wieder beschwerdefrei sein, dann drücke ich damit *meinen* persönlichen Grad an Überzeugung aus – meinen Glauben auf Basis der mir vorliegenden Informationen. Das gilt für quantitative Angaben ebenso wie für qualitative Einschätzungen.

Diese Wahrscheinlichkeiten sind keine objektiven Eigenschaften der Welt – sie existieren nicht unabhängig von der beurteilenden Person. Wie Spiegelhalter treffend bemerkt: Es gibt kein „Wahrscheinlichkeits-Meter“, so wie es einen Tacho-Meter oder ein Thermo-Meter gibt.

David Spiegelhalter, Professor für Statistik an der Universität Cambridge, gelangt zu dem Schluss, dass Wahrscheinlichkeiten im alltäglichen Leben nicht existieren. Dennoch ist es oft nützlich, so zu handeln, als ob sie existierten (2).

Diagnose

Aus diagnostischer Sicht lassen sich folgende Fragen formulieren:

1. Wie entsteht die Schätzung einer 20%igen Wahrscheinlichkeit für eine Lungenentzündung?

Ein 26-jähriger, zuvor gesunder Mann stellt sich mit seit drei Tagen bestehendem Husten (ohne Auswurf) und subjektivem Fiebergefühl am Abend vor.

Anamnestisch ergibt sich zunächst kein starker Hinweis auf eine Lungenentzündung. Aufgrund seines Wissens und seiner Erfahrung schätzt der Arzt die Wahrscheinlichkeit für eine Lungenentzündung auf ca. 10 % – subjektiv als „eher unwahrscheinlich“. Diese 10 % sind Ausdruck seines persönlichen Grades an Überzeugung. Bei der Auskultation hört er vereinzelte Rasselgeräusche. Diese neue Information verändert seine Einschätzung: Er revidiert die Wahrscheinlichkeit auf 20 %.

2. Wie ist die 20%ige Wahrscheinlichkeit zu interpretieren?

Die Wahrscheinlichkeit von 20 % bedeutet nicht, dass der Patient „zu 20 %“ krank ist. Der Patient hat eine Lungenentzündung oder er hat keine. Die Wahrscheinlichkeit von 20 % ist folgendermassen zu interpretieren: Von 100 Patienten mit den gleichen Merkmalen (männlich, etwa 26 Jahre alt, keine anderen Krankheiten, Husten ohne Auswurf seit drei Tagen, Gefühl von abendlich leicht erhöhter Körpertemperatur und auskultatorisch wenig Rasselgeräuschen) haben 20 eine Lungenentzündung und 80 nicht.

3. Welche klinischen Konsequenzen folgen daraus?

Szenario A – abwartendes Vorgehen:

Der Arzt schätzt die Wahrscheinlichkeit als niedrig ein. Da der Patient jung, ohne Komorbiditäten und in gutem Allgemeinzustand ist, entscheidet er sich gegen weitere Diagnostik und gegen eine Antibiotikatherapie. Er empfiehlt genügend Tee zu trinken und sich zu schonen und wenn die

Beschwerden in den nächsten Tagen stärker werden soll er wieder in die Praxis kommen.

Szenario B – diagnostische Absicherung:

Ein anderer Arzt empfindet die Unsicherheit als zu hoch. Er veranlasst eine CRP-Bestimmung (leicht erhöht; das spricht für eine Entzündung im Körper) und ein Thoraxröntgenbild (unauffällig). Aufgrund des negativen Bildbefunds verzichtet auch er auf ein Antibiotikum und empfiehlt genügend Tee zu trinken und sich zu schonen.

*Szenario C: **84-jährige Frau** mit Diabetes mellitus, Herzinsuffizienz (beides behandelt) und möglicher Lungenentzündung (Husten seit drei Tagen und geringgradig erhöhte Körpertemperatur). Die geschätzte Wahrscheinlichkeit beträgt ebenfalls 20 %. Da eine Lungenentzündung bei dieser älteren Frau im Vergleich zu einer Lungenentzündung bei einer sonst gesunden und jungen Person «gefährlicher» (Atemversagen, Dekompensation der Herzinsuffizienz, Sepsis u.a.) ist, ist eine Blutuntersuchung (Entzündungsparameter) und ein Thoraxröntgenbild dringend empfohlen. Eine Lungenentzündung ist «auszuschliessen», oder wenn die Patientin eine Lungenentzündung hat mit einem Antibiotikum zu behandeln.*

Die Einschätzung von Wahrscheinlichkeiten in der klinischen Praxis orientiert sich an typischen Fallkonstellationen. Wichtig ist es, diese Wahrscheinlichkeiten zu interpretieren – zum Beispiel beim 26-jährigen Patienten keine weiteren Untersuchungen zu veranlassen, bei der 84-jährigen Patientin allerdings schon. Diese Interpretationen können unterschiedlich ausfallen und unterschiedliche Ärzte trotz gleicher Ausgangsdaten zu unterschiedlichen Entscheidungen, zum Beispiel Röntgenbild ja oder nein, kommen.

Krankheitswahrscheinlichkeiten aufgrund der Anamnese und den Befunden der körperlichen Untersuchung zu schätzen ist oft herausfordernd. Welchen Einfluss die Entwicklungen im Bereich der «Künstlichen Intelligenz (KI)» auf den diagnostischen Prozess und die Schätzung von Wahrscheinlichkeiten haben werden, ist noch nicht

klar. Es ist aber sehr wahrscheinlich, dass es in Zukunft Algorithmen geben wird, die in der Lage sind, Wahrscheinlichkeiten von Krankheiten zu schätzen. Es gibt heute schon Algorithmen, die in der Diagnosestellung, vor allem wenn es dabei um Bilderkennung geht, gleich gut oder sogar besser sind als erfahrene Medizinerinnen und Mediziner (3-6).

Prognose

Aus prognostischer Perspektive stellt sich folgende Frage: Welche Auswirkungen hat Behandlung A im Vergleich zu keiner Behandlung oder zu den Alternativen B oder C?

Zur Veranschaulichung zwei Beispiele: Eine 30-jährige Frau leidet an einer peripheren Fazialisparese (Gesichtslähmung), ohne weitere Begleiterkrankungen. Studien zeigen, dass eine kurzzeitige Behandlung (10 Tage) mit Prednison (einem Steroid) gegenüber einem Placebo wirksam ist. Die Wirksamkeit lässt sich an verschiedenen Parametern messen – für die Patientin ist vor allem relevant, wie schnell die Gesichtslähmung wieder vollständig verschwindet.

Mehrere Studien haben diesen Aspekt untersucht; ihre Ergebnisse wurden in einem systematischen Review mit Metaanalyse zusammengefasst (7). Die Resultate der einzelnen Studien ähneln sich. Exemplarisch sei hier die Lancet-Studie aus dem Jahr 2008 genannt (8):

Nach drei Monaten war die vollständige Wiederherstellung der Nervenfunktion bei 60 % der Patienten in der Prednison-Gruppe erreicht – verglichen mit 47 % in der Placebo-Gruppe. Nach einem Jahr lagen die Werte bei 62 % versus 50 %. Das Fazit: Zwar heilt die Fazialisparese auch ohne Therapie bei vielen Patienten ab, doch mit Prednison ist die Wahrscheinlichkeit einer vollständigen Genesung deutlich höher.

Ein nächster wichtiger Aspekt ist die Frage nach möglichen Nebenwirkungen. In den Studien wurden Nebenwirkungen wie Kopfschmerzen oder Schwindel in beiden Gruppen – also sowohl bei der Prednison- als auch bei der

Placebogruppe – gleich häufig beobachtet. Daher empfehlen die meisten Ärztinnen und Ärzte eine kurzzeitige Therapie mit Prednison.

Der 30-jährigen Patientin kann die Ärztin erklären, dass ihre Chancen auf eine vollständige Genesung durch die Einnahme von Prednison erhöht werden. Was jedoch nicht vorhergesagt werden kann, ist, ob sie persönlich zu den 60 % gehört, bei denen die Lähmung nach drei Monaten verschwunden ist. Medizinische Prognosen liefern Wahrscheinlichkeiten, aber keine individuellen Gewissheiten.

Beispiel Frau Kunz:

Bei Frau Kunz treten nach längerem Stehen oder Gehen Schmerzen im Gesäss und in beiden Oberschenkeln auf. Wenn sie sich auf einen Stuhl setzt und den Oberkörper nach vorne beugt, verschwinden die Schmerzen in ein oder zwei Minuten. Diese Beschwerden sprechen für einen engen Spinalkanal (Lumbalstenose). Wie die Schmerzen zustande kommen, weiss man nicht so genau. Wenn sich ein Patient durch die Beschwerden eingeschränkt fühlt, kann man den Wirbelkanal chirurgisch etwas weiter machen. Aus Studien ist bekannt, dass etwa 60 % nach einer Operation (Erweiterung des Wirbelkanals) beschwerdefrei sind und 40 % nicht (7). Das sind Mittelwerte von hunderten Patienten, die operiert wurden. Diese hunderten Patienten sind keine homogene Population und damit gibt es möglicherweise Gruppen innerhalb der Population bei denen die Wahrscheinlichkeit, dass sie nach der Operation beschwerdefrei sind vermutlich grösser oder kleiner ist als der durchschnittliche Wert von 60 %. Aus Studien weiss man, dass das «Gemüt» der Patienten eine Rolle spielt, ob jemand nach der Operation beschwerdefrei ist oder nicht. Bei Menschen, die eher depressiven oder pessimistischen Gemütes sind, ist die Wahrscheinlichkeit einer postoperativen Beschwerdefreiheit geringer als bei optimistischen Personen mit einer eher euphorischen Stimmungslage (8). Da Frau Kunz eine Person mit positiver Lebenseinstellung ist, liegt die Wahrscheinlichkeit, dass die Operation erfolgreich ist, bei geschätzten (persönliche Überzeugung) 80 %.

Mehr zum Thema im Artikel «Prognose» (in Vorbereitung).

Zusammenfassung: Die Rolle von Wahrscheinlichkeiten in der Medizin

Wahrscheinlichkeiten spielen in der medizinischen Praxis eine zentrale Rolle, denn medizinisches Wissen ist fast ausschliesslich ein Wissen über Wahrscheinlichkeiten. Diese Wahrscheinlichkeiten geben in der Regel an, wie überzeugt eine Ärztin oder ein Arzt auf Basis der verfügbaren Informationen ist. Diese Einschätzungen bilden die Basis für die Entscheidung, ob weitere Untersuchungen notwendig sind und welche Behandlungen empfohlen werden. Auch wenn Wahrscheinlichkeiten als solche in der Realität nicht „existieren“, sind sie als gedankliches Konzept im medizinischen Alltag ausgesprochen hilfreich.

Literatur

1. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet*. 1994;344(8934):1383–9.
2. Spiegelhalter D. Why probability probably doesn't exist (but it is useful to act like it does). *Nature*. 2024;636(8043):560–3.
3. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature*. 2020;577(7788):89–94.
4. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*. 2016;316(22):2402–10.
5. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115–8.
6. Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med*. 2019;25(1):65–9.
7. Steurer J, Nydegger A, Held U, Brunner F, Hodler J, Porchet F, et al. LumbSten: the lumbar spinal stenosis outcome study. *BMC Musculoskelet Disord*. 2010;11:1–6.
8. 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.

Anhang:

Ein weiteres Thema im Unterricht war das Konzept der *bedingten Wahrscheinlichkeiten*, das sich als besonders herausfordernd erwies. Bedingte Wahrscheinlichkeiten geben an, mit welcher Wahrscheinlichkeit ein Ereignis A eintritt, unter der Bedingung, dass das Ereignis B bereits eingetreten ist.

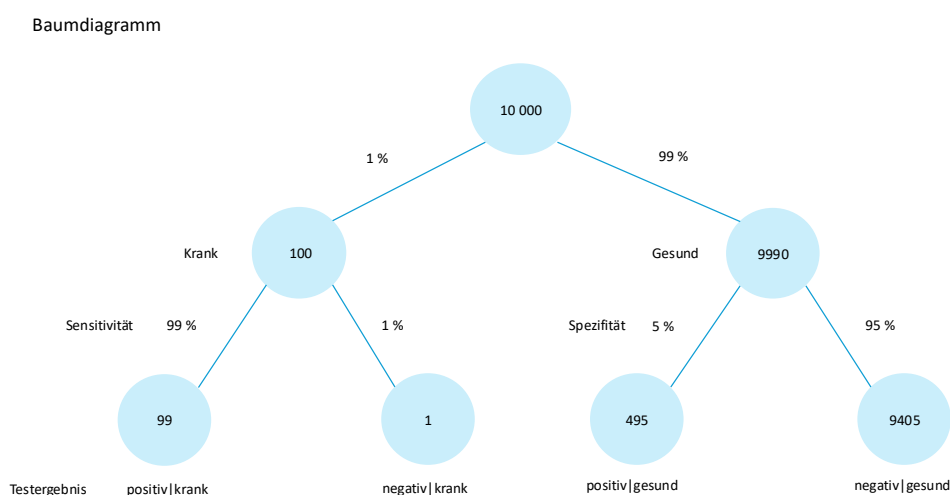
Beispiel

Wie gross ist die Wahrscheinlichkeit, dass eine Person krank ist (A), wenn der Test positiv ist (B). Der Test für eine seltene Krankheit hat eine Sensitivität von 99 %; nur 1 % der Untersuchten hat die Krankheit und der Test ist bei 5 % der Gesunden falsch positiv (Spezifität 95 %).

Diese Frage kann mit der Anwendung der Bayessche Formel berechnet werden, eine einfachere und besser verständlichere Methode (zumindest für mich) ist das Baumdiagramm.

Nehmen wir an, dass 10 Tausend Frauen untersucht werden (z.B. Screening für Brustkrebs).

Wie im Beispiel erwähnt, sind 1 % dieser 10 Tausend krank (also 100) und die Sensitivität der Untersuchung (Mammographie) beträgt 99 % und die Spezifität 95 %. Die Wahrscheinlichkeit bei positivem Testergebnis krank zu sein beträgt 16.6 %.



Die Wahrscheinlichkeit krank zu sein, wenn das Testergebnis positiv ist, ist $99/99 + 495 = 16.6 \%$

Derartige Fragestellungen mögen in der Epidemiologie – also in der bevölkerungsbezogenen Medizin, z.B. Impfen oder Screening-Programme, von Relevanz sein, da in der Epidemiologie das Kollektiv im Fokus steht und nicht das einzelne Individuum. Für den Kliniker, der einen individuellen Patienten behandelt sind derartige Rechnungen – wie im Artikel «Sensitivität und Spezifität ausführlicher beschrieben – nicht von hoher Relevanz.