

## EbM-Splitter 11

### Nutzen einer diagnostischen Tests in der Praxis: prädiktive Werte

In den beiden letzten EbM-Splittern [6, 7] wurden die Maßzahlen Sensitivität (Wahrscheinlichkeit, eine kranke Person als krank zu erkennen) und Spezifität (Wahrscheinlichkeit, eine gesunde Person als gesund zu erkennen) vorgestellt und deren Nutzen zur Beurteilung der Güte eines diagnostischen Tests diskutiert. Zur Bewertung des Ergebnisses einer diagnostischen Maßnahme im klinischen Alltag, d.h. nach der Anwendung des diagnostischen Tests, sind Sensitivität und Spezifität jedoch nicht geeignet. Im klinischen Alltag liegt vielmehr ein positives oder negatives Testergebnis vor und man ist an der Wahrscheinlichkeit interessiert, dass dieses Testergebnis richtig ist.

Die Wahrscheinlichkeit für die Richtigkeit eines Testergebnisses ist sowohl für den Patienten als auch den behandelnden Arzt von zentraler Bedeutung: Der Patient will verlässlich über seinen Krankheitsstatus informiert werden, während für den behandelnden Arzt die weiteren diagnostischen oder therapeutischen Maßnahmen von der Verlässlichkeit des Testergebnisses abhängen. Im Folgenden wird die Bestimmung der Wahrscheinlichkeit für ein richtig positives bzw. richtig negatives Ergebnis dargestellt: Diese Wahrscheinlichkeit wird auch als prädiktiver Wert (oder Vorhersagewert) bezeichnet.

---

*"Mit Ausnahme weniger Dinge, die ganz an der Oberfläche liegen, beruht alle Diagnostik auf Wahrscheinlichkeitsrechnung."*

Prof. Dr. C. Gerhardt, zur Eröffnung der Klinik für Innere Medizin an der Berliner Universität am 2. November 1885 [1]

---

### Definition der prädiktiven Werte

Im vorletzten EbM-Splitter wurde u.a. die Faseroptiktransillumination zur Diagnostik der Approximalkaries diskutiert [3]. Das Ergebnis dieser diagnostischen Maßnahme ist in Tabelle 1 wiedergegeben; für die Kenngrößen Sensitivität und Spezifität ergeben sich Werte von 67,7% (377/557) und 99,5% (5139/5165), d.h. eine mittlere Sensitivität und eine sehr hohe Spezifität. Die Wahrscheinlichkeit, dass eine Person bei einem positiven Testergebnis tatsächlich krank ist, wird als *positiv prädiktiver Wert* (positiver Vorhersagewert) bezeichnet. Entsprechend wird die Wahrscheinlichkeit, dass eine Person bei einem negativen Testergebnis tatsächlich gesund ist, als *negativ prädiktiver Wert* (negativer Vorhersagewert) bezeichnet.

Für die Daten in Tabelle 1 können diese Kenngrößen direkt bestimmt werden. Der positive Vorhersagewert ist der Anteil der tatsächlich kariösen Zähne (gemäß Ergebnis der röntgenologischen Untersuchung) unter den mittels der faseroptischen Untersuchung als kariös eingestuften Zähnen. Von 403 ( $a+b$ ) als kariös eingestuften Zähnen, sind 377 ( $a$ ) Zähne tatsächlich kariös. Der positiv prädiktive Wert (PPW) berechnet sich somit zu 93,5% ( $a/(a+b) = 377/403$ ). In analoger Weise ist der negativ prädiktive Werts (NPW) definiert als  $d/(c+d)$ . Im Beispiel ergibt sich 96,6% (5139/5319). Wie im nächsten Abschnitt beschrieben wird, können diese geschätzten Werte nicht generell, sondern nur für die Tabelle 1 zugrundeliegende Situation verwendet werden.

		Röntgenologische Untersuchung		S
		Kariös	Gesund	
Faseroptische Untersuchung	Kariös	377 <i>a</i>	26 <i>b</i>	403 <i>a+b</i>
	Gesund	180 <i>c</i>	5139 <i>d</i>	5319 <i>c+d</i>
S		557 <i>a+c</i>	5165 <i>b+d</i>	5722 <i>N</i>

Tab. 1: Vergleich der faseroptischen und röntgenologischen Kariesdiagnostik [3]

### Einfluss der Prävalenz auf die prädiktiven Werte

Im verwendeten Beispiel sind 557 von 5722 Zähnen tatsächlich kariös, was einem Anteil von 9,7% entspricht. Dieser Anteil wird üblicherweise als Prävalenz bezeichnet, d.h. die Wahrscheinlichkeit, dass eine Person aus einer Population die interessierende Krankheit (hier: Karies) hat. Neben der Sensitivität und Spezifität hängen die prädiktiven Werte in erheblichem Maße von der zugrundeliegenden Prävalenz ab.

Allgemeine Berechnungsformeln für den positiv prädiktiven Wert (PPW) und den negativ prädiktiven Wert (NPW) sind durch

$$PPW = \frac{\text{Sensitivität} \times \text{Prävalenz}}{\text{Sensitivität} \times \text{Prävalenz} + (1 - \text{Spezifität}) \times (1 - \text{Prävalenz})}$$

$$NPW = \frac{\text{Spezifität} \times (1 - \text{Prävalenz})}{\text{Spezifität} \times (1 - \text{Prävalenz}) + (1 - \text{Sensitivität}) \times \text{Prävalenz}}$$

gegeben [2].

Die Richtigkeit dieser Formeln kann für die Daten von Tabelle 1 (Sensitivität: 67,7%, Spezifität: 99,5%, Prävalenz: 9,7%) überprüft werden:

$$PPW = \frac{67,7\% \times 9,7\%}{67,7\% \times 9,7\% + (100\% - 99,5\%) \times (100\% - 9,7\%)} = 93,6\%$$

Dies entspricht gerade dem im letzten Abschnitt berechneten positiv prädiktiven Wert.<sup>1</sup>

Für den negativ prädiktiven Wert erhält man

$$NPW = \frac{99,5\% \times (100\% - 9,7\%)}{99,5\% \times (100\% - 9,7\%) + (100\% - 67,7\%) \times 9,7\%} = 96,6\%$$

Dies entspricht exakt dem Ergebnis im letzten Abschnitt.

Die Rechenformeln für die prädiktiven Vorhersagewerte können für beliebige

<sup>1</sup> Der geringe Unterschied (93,5% vs. 93,6%) ergibt sich durch Ungenauigkeiten beim Runden.

Patientenpopulationen, die sich in der Prävalenz unterscheiden, herangezogen werden.

In einer Population mit einer Karies-Prävalenz von lediglich 1% berechnet sich der positive Vorhersagewert zu 57,8% und der negative Vorhersagewert zu 99,7%. In diesem Fall kann der tatsächliche Zahnstatus bei Vorliegen eines positiven Testergebnisses sehr ungenau vorhergesagt werden, da nur knapp mehr als die Hälfte der Vorhersagen (57,8%) richtig sind. Andererseits kann bei Vorliegen eines negativen Testergebnisses die Diagnose als gesichert angesehen werden, da 99,7% der Vorhersagen richtig sind.

Bei einer Karies-Prävalenz von 30% ergeben sich hingegen Werte von 98,3% für den PPW und von 87,8% für den NPW. In diesem Fall liegen somit sowohl bei einem positiven als auch bei einem negativen Testergebnis relativ gute Vorhersagen vor.

In einer Population mit sehr hohem Anteil kariöser Zähne berechnet sich bei einer Prävalenz von 90% der positiv prädiktive Wert zu 99,9% und der negativ prädiktive Wert zu 25,5%. D.h. bei dieser hohen Prävalenz ist ein positives Testergebnis sehr verlässlich, wohingegen ein negatives Testergebnis nur in 1 von 4 Fällen richtig ist.

Als Merkregel lässt sich festhalten (Tab. 2):

- Mit zunehmender Prävalenz (bei festen Werten für Sensitivität und Spezifität) wird der PPW größer und der NPW kleiner.
- Dementsprechend wird mit abnehmender Prävalenz der PPW kleiner und der NPW größer.

Aus diesem Grunde ist in Kollektiven mit einem hohen Anteil erkrankter Personen (hohe Prävalenz) ein positives Testergebnis verlässlicher als in einem Kollektiv mit einem geringen Anteil erkrankter Personen (niedrige Prävalenz).

<i>Sens: 99%; Spez: 95%</i>		
<b>Prävalenz</b>	<b>PPW</b>	<b>NPW</b>
1%	16,67%	99,99%
2%	28,78%	99,98%
3%	37,98%	99,97%
4%	45,21%	99,96%
5%	51,03%	99,94%
6%	55,83%	99,93%
7%	59,84%	99,92%
8%	63,26%	99,91%
9%	66,20%	99,90%
10%	68,75%	99,88%
15%	77,75%	99,81%
20%	83,19%	99,74%
25%	86,84%	99,65%
30%	89,46%	99,55%
35%	91,42%	99,44%
40%	92,96%	99,30%
45%	94,19%	99,15%
50%	95,19%	98,96%

<i>Sens: 70%; Spez: 96%</i>		
<b>Prävalenz</b>	<b>PPW</b>	<b>NPW</b>
1%	15,02%	99,69%
2%	26,32%	99,37%
3%	35,12%	99,04%
4%	42,17%	98,71%
5%	47,95%	98,38%
6%	52,76%	98,04%
7%	56,84%	97,70%
8%	60,34%	97,35%
9%	63,38%	97,00%
10%	66,04%	96,64%
15%	75,54%	94,77%
20%	81,40%	92,75%
25%	85,37%	90,57%
30%	88,24%	88,19%
35%	90,41%	85,60%
40%	92,11%	82,76%
45%	93,47%	79,64%
50%	94,59%	76,19%

Tab. 2: Abhängigkeit der positiven und negativen prädiktive Werte (PPW; NPW) von der Prävalenz der Krankheit

- (a) bei einem „idealen“ diagnostischen Test mit einer Sensitivität von 99% und einer Spezifität von 95%;
- (b) bei einem Test mit einer Sensitivität von 70% und einer Spezifität von 96%, was nach Widmer [8] den Mindestanforderungen für diagnostische Tests zur Diagnostik von kranio-mandibulären Dysfunktionen entspricht.

### **Sichere Diagnosen (SpPin, SnNout)**

In einigen Sonderfällen können klare Aussagen bei Vorliegen eines bestimmten Testergebnisses gemacht werden:

- Bei einer Sensitivität nahe 100% (und einer genügend kleinen Prävalenz) ist der negativ prädiktive Werte nahe 100% (vgl. Tab. 2). Ein diagnostischer Test mit hoher Sensitivität hat demnach die größte Aussagekraft, wenn das Testergebnis negativ ausfällt (sehr hohe Wahrscheinlichkeit der Abwesenheit der Erkrankung). Sackett et al. [4, 5] verwenden für diesen Fall die Abkürzung *SnNout* (*very high Sensitivity: Negative result rules out the diagnosis/disease*).
- Bei einer Spezifität nahe 100% (und einer genügend großen Prävalenz) ist der positiv prädiktive Wert nahe 100%. Sackett et al. [4, 5] verwenden für diesen Fall die Abkürzung *SpPin* (*very high Specificity: Positive result rules in the diagnosis/disease*). Mit anderen Worten: Ein hochspezifischer diagnostischer Test hat die größte Aussagekraft, wenn das Testergebnis positiv ausfällt (deutlicher Hinweis auf das Vorliegen der Erkrankung bzw. Bestätigung der Diagnose). Im letzten Abschnitt wurde beispielsweise gezeigt, dass bei einer Sensitivität von 67,7%, einer Spezifität von 99,5% und einer Prävalenz von 30% der PPW gleich 98,3% ist.

## Hilfsmittel zur Abschätzung der Güte diagnostischer Tests

Die Berechnung von Sensitivität und Spezifität oder (bei bekannter Sensitivität, Spezifität und Prävalenz) von PPW und NPW kann bequem mit Hilfe von Rechnern (Calculatoren) erfolgen, auf die im *Weltweiten Netz* z. B. unter [www.clinical-trials.de/de/Werkzeuge/Online\\_Hilfsmittel/online\\_hilfsmittel.html](http://www.clinical-trials.de/de/Werkzeuge/Online_Hilfsmittel/online_hilfsmittel.html) kostenfrei zugegriffen werden kann.

## Festlegung der Prävalenz für einen einzelnen Patienten (Vortestwahrscheinlichkeit)

In diesem Splitter wurde gezeigt, dass die objektiven Eigenschaften eines diagnostischen Tests (Sensitivität und Spezifität) für die Bewertung eines Testergebnisses **nicht** ausreichend sind, sondern dass die Prävalenz einen entscheidenden Einfluss auf die patientenbezogenen Aussagen hat. Das Problem in der Praxis besteht darin, für einen Patienten eine (verlässliche) Angabe zur Prävalenz zu finden. Diese Angabe existiert aber oftmals nicht, so dass eine subjektive Einschätzung durch den Arzt notwendig ist. Diese Einschätzung sollte auf der Anamnese des Patienten und gegebenenfalls auf weiteren Befunden beruhen. Hierbei kann eine grobe Einteilung in Risikogruppen hilfreich sein (niedrige, mittlere, hohe Wahrscheinlichkeit für das Vorliegen der Erkrankung). Die Einschätzung sollte vom Arzt unabhängig vom Ergebnis des diagnostischen Test, d.h. **vor** der Durchführung des Tests, festgelegt werden. Aus diesem Grunde wird anstelle von „Prävalenz“ zunehmend der Begriff „Vortestwahrscheinlichkeit“ verwendet.

## Literatur

1. Gerhardt, C.: Ueber Diagnostik und Therapie. Dtsch Med Wochenschr 11, 778 (1885).
2. Gross, R., Löffler, M.: Prinzipien der Medizin. Springer, Berlin 1997, 289-290.
3. Heinrich, R., Künzel, W., Tawfiq, H.: Approximale Kariesdiagnostik – Vergleich klinischer, faseroptischer und röntgenographischer Diagnoseverfahren (Approximalkariesdiagnostik). Dtsch Zahn Mund Kieferheilkd Zentralbl 79, 535 (1991).
4. Sackett, D.L., Haynes, R.B., Guyatt, G.H., Tugwell, P.: Clinical Epidemiology. A Basic Science for Clinical Medicine. Little, Brown and Company, Boston 1991, 83.
5. Sackett, D.L., Straus, S.E., Richardson, W.S., Rosenberg, W., Haynes, R.B.: Evidence-Based Medicine. How to Practice and Teach EBM. 2nd ed. Churchill Livingstone, Edinburgh 2000, 249-250.
6. Schwarzer, G., Türp, J.C., Antes, G.: EbM-Splitter: Die Vierfeldertafel (in Diagnosestudien): Sensitivität und Spezifität. Dtsch Zahnärztl Z 57, 333 (2002).
7. Schwarzer, G., Türp, J.C., Antes, G.: EbM-Splitter: Sensitivität und Spezifität: Auswirkung der Wahl des Trennpunktes. Dtsch Zahnärztl Z 2002;57:573-575.
8. Widmer, C.G.: Review of the literature. B: Reliability and validation of examination methods. J Craniomand Disord Facial Oral Pain 6, 318 (1992).

Guido Schwarzer, Freiburg  
Jens C. Türp, Freiburg / Basel  
Gerd Antes, Freiburg