

Zur Notwendigkeit einer abgestuften Diagnostik in der Allgemeinmedizin als Konsequenz des Bayes'schen Theorems

Antonius Schneider¹, Geert-Jan Dinant² und Joachim Szecsenyi¹

¹ Abteilung Allgemeinmedizin und Versorgungsforschung, Universitätsklinikum Heidelberg

² Department of General Practice, Faculty of Medicine, University Maastricht, Niederlande

Zusammenfassung

Kernproblem der medizinischen Diagnostik ist die Abhängigkeit der prädiktiven Werte diagnostischer Tests von der Prävalenz von Erkrankungen. Die Zusammenhänge von Vortest- und Nachtestwahrscheinlichkeit lassen sich durch das Bayes'sche Theorem gut beschreiben. Durch Modifikationen des Theorems lassen sich auch Zusammenhänge von Prävalenz und der Wahrscheinlichkeit für eine Fehldiagnose aufzeigen. Demzufolge ist im niedrigprävalenten Bereich die Wahrscheinlichkeit geringer, dass ein Testpositiver krank ist (PPV) und die Wahrscheinlichkeit höher, dass ein Testpositiver gesund, also falsch positiv, ist (FPPV). Für die Höhe von PPV und FPPV ist vor allem die Spezifität eines Tests maßgeblich. Dafür ist im niedrigprävalenten Bereich grundsätzlich die Wahrscheinlichkeit höher, dass ein Testnegativer auch tatsächlich gesund ist (NPV) und die Wahrscheinlichkeit geringer, dass ein Testnegativer krank, also falsch negativ, ist (FNPV). Die Höhe von NPV und FNPV variieren im Niedrigprävalenzbereich nur geringfügig in Abhängigkeit von Sensitivität und Spezifität. Diese statistischen Zusammenhänge können die typische Arbeitsweise in der allgemeinmedizinischen Praxis erklären, in der mit einer unausgelesenen Patientenklinte gearbeitet wird. Um die Nachtestwahrscheinlichkeit (PPV) zu erhöhen und FPPV zu vermindern, muss sich der Hausarzt einer Kombination von klinischer Erfahrung, dem Faktor Zeit und abgestufter apparativer Diagnostik bedienen. Es besteht ein dringender Bedarf an diagnostischen Studien in der Primärversorgung, um diagnostische Abläufe in der Patientenversorgung verbessern zu können.

Sachwörter: Allgemeinmedizin – Bayes'sches Theorem – Sensitivität – Spezifität – Prävalenz – Vorhersagewahrscheinlichkeiten

Stepwise diagnostic workup in general practice as a consequence of the Bayesian reasoning

Summary

The dependency of the predictive values of tests on the prevalence of diseases is an escrow issue of clinical diagnostics. The relation between pre-test probability and post-test probability is well explained by Bayes' theorem, and the relation between prevalence and false diagnoses can be described well by modifying this theorem. In cases of low prevalence the positive predictive value (PPV) is lower and the false-positive predictive value (FPPV) higher. These aspects mainly depend on the test specificity. But basically, in cases of low prevalence there is a higher negative predictive value (NPV) and a lower false negative predictive value (FNPV). Depending on the sensitivity and specificity, NPV and FNPV vary only slightly in low prevalence ranges. These statistical relations are able to explain the typical mode of operation of general practitioners with their unselected patients. In order to increase PPV and decrease FPPV in the diagnostic workup, the GP must use his clinical experience, time and stepwise diagnostic procedures. More diagnostic studies are necessary to improve the diagnostic workup in patient care.

Key words: General practice – Bayes' theorem – Sensitivity – Specificity – Prevalence – Predictive values

Einleitung

Zahlreiche, häufig von spezialisierten Abteilungen ausgehende Studien, weisen auf Defizite in der Diagnostik verschiedener Erkrankungen im primärärztlichen Bereich hin. Entsprechende Studien aus dem Bereich der Versorgungsforschung zeigen dies beispielsweise bei der Erkennung von obstruktiven Atemwegserkrankungen [1] oder Arterieller Hypertonie [2]. Anführen lassen sich auch psychosomatisch/psychiatrische Studien zur Erkennung von Depressionen oder generell psychischer Komorbidität [3, 4]. Hierdurch kam es wiederholt zu Kontroversen über die Zuverlässigkeit der Arbeitsweise von Hausärzten¹ [5]. Dabei könnten Stärken und Schwächen der Primärversorgung besser eingeschätzt werden, wenn die hausärztliche Arbeitsweise unter epidemiologischen bzw. statistischen Gesichtspunkten untersucht wird.

Das Kernproblem der Treffsicherheit medizinischer Diagnostik betrifft die Tatsache, dass die Wahrscheinlichkeit von Diagnosestellung bzw. -ausschluss und Prävalenz einer Erkrankung miteinander verknüpft sind. Das heißt, die Ereignisse „Jemand mit positivem Testergebnis ist tatsächlich krank“ und „Häufigkeit der Erkrankung in der Bevölkerung“ sind voneinander abhängi-

¹ Zur besseren Lesbarkeit wird die männliche Form verwendet, gemeint sind selbstverständlich beide Geschlechter.

ge Ereignisse. Die Prävalenz entspricht der sogenannten Vortestwahrscheinlichkeit, falls das Patientenkollektiv unausgelesen ist. Bei entsprechender Auslese, z. B. bei Abklärungsdiagnostik bei länger bestehenden Symptomen, erhöht sich diese Vortestwahrscheinlichkeit. Als Konsequenz davon erhöht sich automatisch die Nachtestwahrscheinlichkeit bei gleichbleibender Testgüte. Durch diese Abhängigkeit der Nachtestwahrscheinlichkeit von der Vortestwahrscheinlichkeit kommt es zu erheblichen Verschiebungen der diagnostischen Treffsicherheit, wie bereits Lusted erkannt hatte [6]. Er war zusammen mit Ledley der erste, der die Bedeutung der Bayes'schen Formel, in dem der statistische Zusammenhang zweier voneinander abhängiger Ereignisse ausgedrückt wird, für die medizinische Diagnostik in einem wegweisenden Artikel herausgearbeitet hat [7]. In der vorliegenden Arbeit wird die medizinische Diagnostik im Niedrigprävalenzbereich unter Berücksichtigung des Bayes'schen Theorems reflektiert.

Positive Vorhersagewerte

Diagnosestellung: Wahrscheinlichkeit, dass ein Testpositiver tatsächlich krank ist (PPV)

Sensitivität und Spezifität sind elementare Testparameter, die häufig zur Ein-

schätzung der Testgüte eingesetzt werden. Es handelt es sich bei diesen um epidemiologische Kenngrößen, quasi auf der „Makroebene“, die in größeren Studien ermittelt werden. Für die ärztliche Entscheidung auf der „Mikroebene“ ist jedoch von Bedeutung, wie hoch die Wahrscheinlichkeit ist, dass ein Testpositiver auch tatsächlich krank ist bzw. ein Testnegativer auch wirklich gesund ist. Diese Wahrscheinlichkeiten ergeben sich durch die sogenannten Vorhersagewerte (engl: predictive values). Unter dem positiven Vorhersagewert (positive predictive value = PPV), der häufig auch als Nachtestwahrscheinlichkeit bezeichnet wird, versteht man das Verhältnis von allen „wahr Testpositiven“ zu „allen Testpositiven“. Der Zusammenhang von PPV und Vortestwahrscheinlichkeit wird durch den Satz von Bayes ausgedrückt (sens = Sensitivität, spez = Spezifität, p = Vortestwahrscheinlichkeit)

$$PPV = \frac{sens \cdot p}{sens \cdot p + (1 - spez) \cdot (1 - p)}$$

Anhand dieser Formel lässt sich bereits ersehen, dass PPV bei steigender Vortestwahrscheinlichkeit p zunehmen muss. Beispielsweise hat ein Test mit gegebener Sensitivität von 99% und Spezifität von 98% bei einer Krankheitsprävalenz von 0,1% einen posi-

ven Vorhersagewert (PPV) von 4,7%. Liegt die Prävalenz (= Vortestwahrscheinlichkeit) bei 10%, steigt PPV auf 83,2%. In Abbildung 1 werden diese Zusammenhänge zur leichteren Interpretation graphisch dargestellt. Aus der Formel geht auch hervor, dass ein Test wertlos ist, wenn Sensitivität + Spezifität = 1 sind, denn in diesem Fall gilt dann Nachtestwahrscheinlichkeit = Vortestwahrscheinlichkeit.

Fehldiagnostik bei Gesunden: Wahrscheinlichkeit, dass ein Testpositiver gesund ist (FPPV)

Von entscheidender Bedeutung ist auch die Frage, wie hoch die Wahrscheinlichkeit ist, jemanden fälschlicherweise als krank zu diagnostizieren. Dies zieht negative Konsequenzen nach sich wie beispielsweise Verletzung durch invasive Diagnostik und zusätzliche Krankenhausaufenthalte. Es gilt $FPPV = 1 - PPV$. Der Zusammenhang von FPPV, Vortestwahrscheinlichkeit, Sensitivität und Spezifität wird in folgender Formel ausgedrückt:

$$FPPV = \frac{(1 - spez) \cdot (1 - p)}{sens \cdot p + (1 - spez) \cdot (1 - p)}$$

Die nachfolgende Graphik (Abb. 1) zu PPV und FPPV verdeutlicht die Zu-

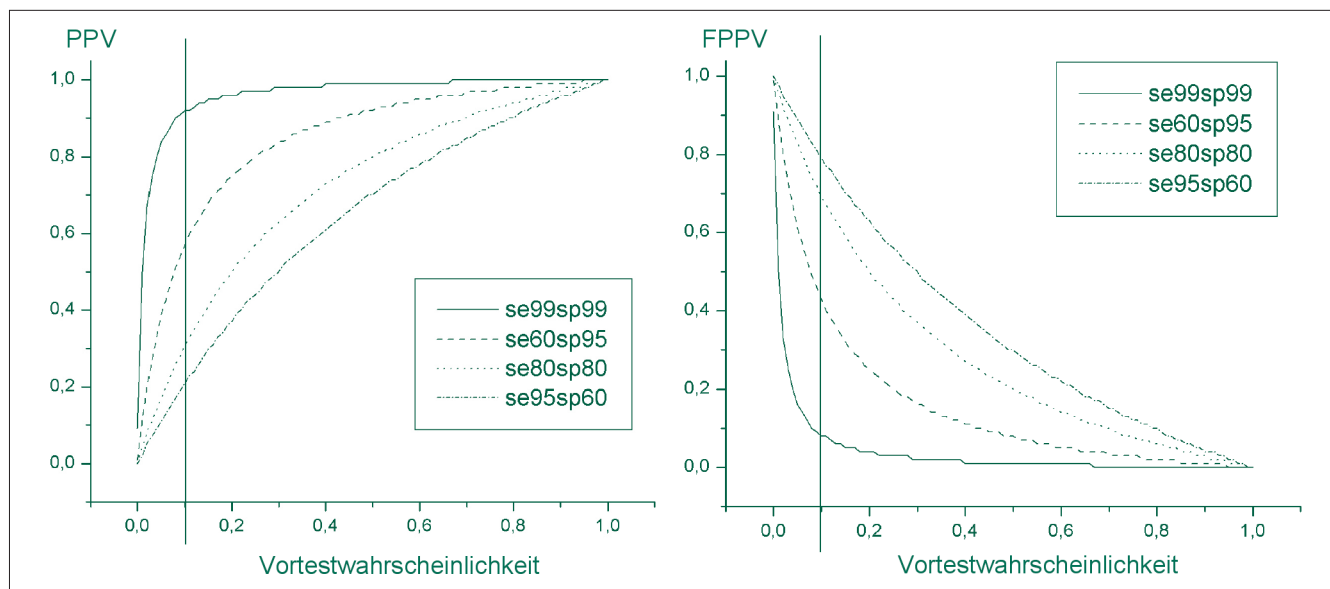


Abb. 1. PPV und FPPV in Abhängigkeit von der Vortestwahrscheinlichkeit, Sensitivität (se) und Spezifität (sp) (z. B. se60sp95 = Sensitivität 60%, Spezifität 95%).

sammenhänge zur Vortestwahrscheinlichkeit unter Berücksichtigung von Sensitivität und Spezifität.

In der Graphik zu PPV lässt sich erkennen, dass insbesondere im Niedrigprävalenzbereich² die Spezifität eine große Rolle spielt, die Sensitivität ist nicht von so hoher Bedeutung. Beispielsweise ist PPV bei einem Test mit einer Spezifität von 60% und Sensitivität von 95% erheblich geringer als bei einer Sensitivität von 60% und Spezifität von 95%. Ein hochspezifischer Test erhöht somit die Wahrscheinlichkeit deutlich, dass ein Testpositiver als krank zu werten ist und dementsprechend einer eingehenderen Untersuchung zugeführt werden muss (engl: specific, positiv, in = spin). Anhand der Graphik zu FPPV lässt sich erkennen, dass insbesondere im Niedrigprävalenzbereich die Wahrscheinlichkeit relativ hoch ist, falsch positiv diagnostiziert zu werden. FPPV hängt hier ebenfalls am stärksten von der Spezifität ab. D.h., nur mit einem sehr spezifischen Test kann die Wahrscheinlichkeit, irrtümlich falsch positiv diagnostiziert zu werden, gering gehalten werden. Die Sensitivität ist hier von nicht so hoher Bedeutung.

² Zur besseren Orientierung innerhalb der Graphik wird bei Vortestwahrscheinlichkeit = 10% eine Markierung gesetzt.

Negative Vorhersagewerte

Krankheitsausschluss: Wahrscheinlichkeit, dass ein Testnegativer tatsächlich gesund ist

Für die Testgüte ist jedoch auch bedeutsam, dass ein Gesunder auch als gesund erkannt wird. Diese Wahrscheinlichkeit wird als negativer Vorhersagewert (negative predictive value = NPV) bezeichnet. Zwischen der Vortestwahrscheinlichkeit und NPV besteht folgender Zusammenhang:

$$NPV = \frac{spez(1-p)}{(1-sens)p + spez(1-p)}$$

Fehldiagnostik bei Kranken: Wahrscheinlichkeit, dass ein Testnegativer krank ist

Von Bedeutung ist jedoch auch, mit welcher Wahrscheinlichkeit Diagnosen übersehen werden können. Somit lohnt sich auch hier eine genauere Analyse. Die Wahrscheinlichkeit, dass ein Testnegativer krank ist (false negative predictive value = FNPV) lässt sich aus dem Verhältnis testnegative Kranke zu allen Testnegativen ermitteln. Es gilt $FNPV = 1 - NPV$, und es zeigt sich folgender Zu-

sammenhang zur Vortestwahrscheinlichkeit:

$$FNPV = \frac{(1-sens)p}{(1-sens)p + spez(1-p)}$$

Die graphische Darstellung erlaubt wiederum eine einfachere Interpretation der Zusammenhänge von negativen Vorhersagewerten, Vortestwahrscheinlichkeit, Sensitivität und Spezifität (Abb. 2).

Die Graphik zu NPV verdeutlicht, dass im Niedrigprävalenzbereich kaum eine Variabilität bezüglich der Nachtestwahrscheinlichkeit besteht, am ehesten ist hier die Sensitivität von Bedeutung. Ein hoch sensibler Test erhöht die Wahrscheinlichkeit, dass ein Testnegativer als gesund zu werten ist. D.h., bei einem hochsensitiven Test ist die Wahrscheinlichkeit, dass ein Testnegativer auch gesund ist, hoch. Der Test ist also geeignet, Krankheit auszuschließen (engl: sensitive, negative, out = snout). In der Graphik zu FNPV fällt auf, dass im Niedrigprävalenzbereich – im Gegensatz zum Hochprävalenzbereich – die Wahrscheinlichkeit, dass ein Kranker fälschlicherweise als negativ diagnostiziert wird, verhältnismäßig gering ist. Dieser Zusammenhang ist ebenfalls relativ unabhängig von der Sensitivität und Spezifität eines Tests.

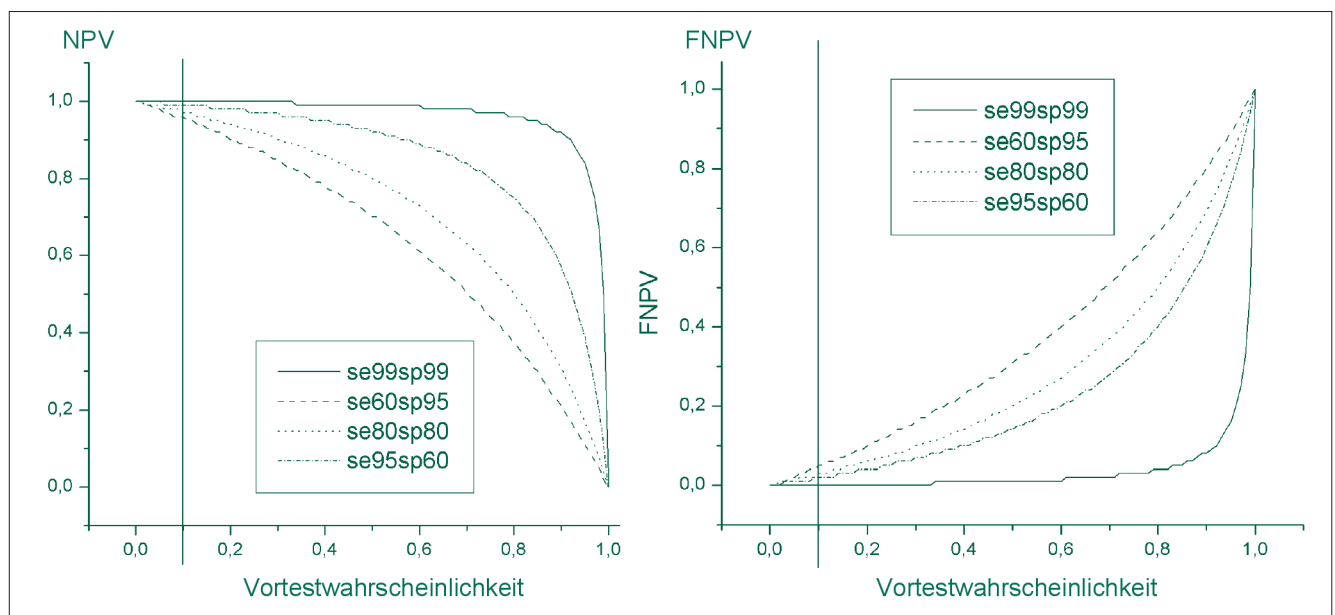


Abb. 2. NPV und FNPV in Abhängigkeit von der Vortestwahrscheinlichkeit, Sensitivität (se) und Spezifität (sp) (z. B. se60sp95 = Sensitivität 60%, Spezifität 95%).

Im Hochprävalenzbereich wird FNPV im Wesentlichen durch die Sensitivität moduliert. D.h., hier ist bei einem hochsensitiven Test die Wahrscheinlichkeit geringer, fälschlicherweise positiv diagnostiziert zu werden.

Prädiktive Werte am Beispiel „Diagnostik bei obstruktiven Atemwegserkrankungen“

Die praktische Bedeutung der oben gezeigten theoretischen Zusammenhänge lässt sich am besten an einem klinischen Beispiel erläutern. Die Versorgungsqualität in Bezug auf Diagnostik und Therapie bei Asthma bronchiale wurde nicht zuletzt im Gutachten des Sachverständigenrates für die Konzer-

tierte Aktion im Gesundheitswesen 2000 in Frage gestellt. Problematisch sind insbesondere diagnostische Unsicherheiten, die auf willkürliche Grenzwerteinteilungen, unterschiedliche Teststrategien und nicht zuletzt auf die variierende Prävalenz innerhalb der unterschiedlichen Gesundheitssektoren zurückzuführen sind. Aus der Sicht eines an einer Klinik tätigen Spezialisten, der mit umfassender Diagnostik ein gefiltertes Patientenkollektiv behandelt, muss die hausärztliche Diagnostik zwangsläufig als unzuverlässig erscheinen, so dass von spezialisierter Seite häufig eine Intensivierung der Diagnostik gefordert wird. Betrachtet man die Testgüte der verschiedenen diagnostischen Tests unter Berücksichtigung der zugrundegelegten Stichprobe, so sieht man jedoch deutliche Variationen der prädiktiven Werte in Abhängigkeit

von der Vortestwahrscheinlichkeit (Prävalenz) der Erkrankung im jeweiligen Kollektiv. In Tabelle 1 ist beispielhaft eine Auswahl von verschiedenen diagnostischen Studien zu obstruktiven Atemwegserkrankungen zusammengestellt.

Obwohl diese diagnostischen Studien im Design variieren und sich auf unterschiedliche Kollektive beziehen, bildet sich das Ausmaß an diagnostischer Unsicherheit bei niedrigerer Krankheitsprävalenz deutlich ab. Beispielsweise besitzt der anamnestic Hinweis auf die Selbstwahrnehmung von Giemen im Hochprävalenzbereich einen verhältnismäßig hohen PPV, damit geht ein niedriger FPPV einher. Im Niedrigprävalenzbereich würde man jedoch zu 97% fehlgeleitet (FPPV), wenn man dieses Zeichen als Leitsymptom verwendet. Die Auskultation der Lunge auf ein ver-

Tabelle 1. Zusammenhänge von Prävalenz, Sensitivität, Spezifität und prädiktiven Werten am Beispiel „Obstruktive Atemwegserkrankungen“.

Diagnostik	Vortestwahrscheinlichkeit	Sens	Spez	PPV	FPPV	NPV	FNPV
Anamnestisch Giemen (Asthma bronchiale und COPD)	Prävalenz 44% (Holleman et al. [8]) N = 164; Überweisung für präoperative Diagnostik an Universitätsklinik	51	84	71	29	69	31
	Prävalenz 1,4% (Dodge et al. [9]) N = 3432; Epidemiologische Studie	57	70	3	97	98	2
Auskultation (Asthma bronchiale und COPD)	Prävalenz 44% (Holleman et al. [8]) N = 164; Überweisung für präoperative Diagnostik an Universitätsklinik	12	99	91	8	60	40
	Prävalenz 21% (Straus et al. [10]) N = 233; Primärversorgung und Überweisungspraxen in 14 Ländern; Patienten mit unbekannter Diagnose (Prävalenz 21%)	32	85	35	65	83	17
PEF-Variabilität (Asthma bronchiale)	Prävalenz 63% (Hunter et al. [11]) N = 110; 69 Patienten mit bekannten Asthma, 20 Patienten mit Pseudoasthma, 21 Patienten als Kontrollgruppe ohne Symptome	43	75	86	14	28	72
	Prävalenz 40% (Thiadens et al. [12]) N = 182; Primärversorgung; Patienten mit persistierendem Husten	36	82	58	42	65	35
Bronchoprovokation (Asthma bronchiale)	Prävalenz 63% (Hunter et al. [11]) N = 110; 69 Patienten mit bekannten Asthma, 20 Patienten mit Pseudoasthma, 21 Patienten als Kontrollgruppe ohne Symptome (Prävalenz 63%)	91	90	97	3	75	25
	Prävalenz 31% (Remes et al. [13]) 1633 Kinder befragt, bei n = 212 Bronchoprovokation durchgeführt; 65 Kinder mit Giemen/Husten; 147 Kinder als Kontrollgruppe ohne Symptome	47	97	89	11	81	19

längertes Expirium mit Giemen liefert im Hochprävalenzkollektiv einen erheblich besseren PPV als im Niedrigprävalenzbereich, obwohl die Spezifität bei letzteren nur wenig schlechter ist. Dafür ist im Niedrigprävalenzkollektiv der negative Vorhersagewert NPV deutlich besser. Der Sachverhalt für die PEF-Variabilität ist ähnlich, in Bezug auf den NPV aber noch prägnanter. Bei der Bronchoprovokation zeigt sich im Niedrigprävalenzbereich eine deutlich geringere Sensitivität, aufgrund der geringeren Prävalenz ist der NPV dennoch höher als im Hochprävalenzbereich. Dafür ist der positive Vorhersagewert im Niedrigprävalenzbereich schlechter, obwohl die Spezifität höher liegt.

Anhand dieser Beispiele wird deutlich, dass Sensitivität, Spezifität und insbesondere die Vorhersagewerte keine a priori feststehenden Eigenschaften des diagnostischen Tests an sich sind, sondern sich immer auf die Anwendung des diagnostischen Tests in einer selektierten Population beziehen [14]. Somit kommt es immer auf den richtigen Zeitpunkt der Selektion an, zu dem die unterschiedlichen diagnostischen Methoden eingesetzt werden. Nur bei passender Vortestwahrscheinlichkeit kann die Diagnostik effizient, d. h. PPV und NPV ausreichend hoch und FPPV und FNPV ausreichend niedrig, werden. Im Folgenden wird diskutiert, wie sich diese Zusammenhänge auf die Diagnostik in den verschiedenen Gesundheitssektoren, aber insbesondere in der Allgemeinmedizin auswirken.

Diskussion

In einer großen epidemiologischen Studie von White et al. [15], die von Green et al. vierzig Jahre später wiederholt wurde [16], konnte zeitüberdauernd nachgewiesen werden, dass innerhalb eines Monats von 1000 Menschen insgesamt 750 über Beschwerden klagen, aber nur 250 Patienten einen Arzt aufsuchen, 9 in ein Krankenhaus eingewiesen werden und nur einer in einer Universitätsklinik aufgenommen wird. Hierbei wird die Selektion deutlich, bei der die Vortestwahrscheinlichkeit von

Erkrankungen von Hausarztpraxis via Facharzt bis zum Universitätskrankenhaus deutlich ansteigt. Aufgrund dieser Variation von Vortestwahrscheinlichkeiten verändern sich die prädiktiven Werte von Tests innerhalb der verschiedenen Gesundheitssektoren. Das hausärztliche Arbeitsfeld entspricht dabei dem sogenannten Niedrigprävalenzbereich.

Die Bedeutung der niedrigen Prävalenz für die Nachtestwahrscheinlichkeit PPV wurde bereits von Gerlach herausgearbeitet, jedoch wurden die Implikationen für den negativen Vorhersagewert (NPV) und das Risiko einer Fehldiagnostik (FPPV und FNPV) dort nicht diskutiert [17]. Hier zeigen sich ebenfalls wichtige Zusammenhänge, denn als Folge dieser Niedrigprävalenz ist in der Allgemeinmedizin die Wahrscheinlichkeit hoch, dass ein Testnegativer auch tatsächlich gesund ist (NPV). Zugleich ist die Wahrscheinlichkeit gering, dass ein Kranker fälschlicherweise testnegativ ist (FNPV). Diese Sachverhalte ergeben sich im Niedrigprävalenzbereich relativ unabhängig von Sensitivität und Spezifität.

Dafür besteht in der Allgemeinmedizin eine geringere Wahrscheinlichkeit, dass ein Testpositiver krank ist (PPV). Hier wird die Problematik von Screeningverfahren offensichtlich. Ein hochsensitiver Test ist hier nutzlos, wenn er nicht zudem eine ausreichende Spezifität bietet. Denn nur über eine hohe Spezifität ist eine aussagefähige PPV zu erreichen. Außerdem besteht in der Allgemeinmedizin ein hohes Risiko, falsch positiv zu diagnostizieren (FPPV). Auch hier ist die Spezifität von höherer Bedeutung als die Sensitivität. Allerdings muss auch bei hoher Spezifität ein vernünftiges Maß an Sensitivität erreicht werden – orientierend mindestens 30–50%, je nach Höhe der Spezifität – damit ein Test sinnvoll bleibt [18].

Anforderungen an die hausärztliche Diagnostik

Die prävalenzabhängige Variation der Testgüte lässt die Schlussfolgerung zu, dass es aufgrund des epidemiologischen, unselektionierten Patientenlientels in der Allgemeinmedizin er-

heblich leichter ist, einen Gesunden als gesund zu diagnostizieren als in der Klinik. Dort ist es einfacher, einen Kranken als krank zu diagnostizieren, da durch die Filterung die Vortestwahrscheinlichkeit deutlich höher ist. Bei gegebener Sensitivität und Spezifität eines Tests sind PPV und FNPV in der Klinik höher, NPV und FPPV dagegen in der allgemeinmedizinischen Praxis höher. Bezüglich der diagnostischen Tests sollte dementsprechend in der Allgemeinmedizin insbesondere auf die Spezifität geachtet werden, damit die PPV ausreichend hoch und die FPPV gering ist. Im Screening könnten NPV und FNPV kaum beeinflusst werden. Einschränkung muss angemerkt werden, dass sich dieser Zusammenhang nicht auf Situationen mit hoher Vortestwahrscheinlichkeit bezieht. Beispielsweise ist bei Brustschmerzen die Vortestwahrscheinlichkeit für einen Herzinfarkt bei einem rauchenden Diabetiker erheblich höher als für einen jungen Sportler. In ersterem Fall sind selbstverständlich sofort hochsensitive Tests, wie die Bestimmung von TNT³, einzuleiten, bzw. je nach Situation sollte gleich ein Krankentransport veranlasst werden. Im letzteren Fall ist am ehesten von einem Trauma oder einer Gelenksblockade auszugehen.

Um die Nachtestwahrscheinlichkeit (PPV) zu erhöhen und FPPV zu vermindern ist es sinnvoll, Patienten im Vorfeld zu selektionieren, um die Vortestwahrscheinlichkeit zu erhöhen. Dies geschieht durch eine präzise Anamnese (Fragen an den Patienten als diagnostischer Test) und korrekte Interpretation klinischer Zeichen (körperliche Untersuchung als weiterer diagnostischer Test). Die klinische Erfahrung erlaubt hier eine sukzessive Erhöhung der Vortestwahrscheinlichkeit. Diese wird auch dadurch erhöht, dass Hausärzte ihre Patienten meist schon längere Zeit kennen und deren Symptome und Befunde der körperlichen Untersuchung besser in einen Verlauf einordnen können. Manchmal wird dies auch als „erlebte

³ TNT = Troponin T, ein serologischer Marker, der auch als Schnelltest das Vorliegen eines Herzinfarktes mit hoher Sicherheit anzeigt oder ausschließt (Sensitivität 91%, Spezifität 92%, 4 h nach Beschwerdebeginn [19]).

Anamnese“ bezeichnet [20]. Aber auch die Zeit ist ein wertvolles diagnostisches Mittel, denn durch „abwartendes Offenlassen“ (engl. watchful waiting) wird die Vortestwahrscheinlichkeit erhöht. Harmlose Symptome vergehen von alleine – als Vorboten einer schwereren Erkrankung dauert es zwangsläufig eine gewisse Zeit, bis eine Erkrankung klinisch offensichtlich wird („test of time“).

Notwendigkeit einer abgestuften Diagnostik

Die typische diagnostische Arbeitsweise des Hausarztes erklärt sich durch eine Kombination von klinischer Erfahrung, „test of time“ und apparativer Diagnostik. Eine zeitlich und technisch abgestufte Diagnostik erlaubt einer Erkrankung im Idealfall eine Spontanheilung, im ernsteren Fall sollte der abwendbare gefährliche Verlauf rechtzeitig erfasst werden. Es macht demnach keinen Sinn, sofort bei minimalen Symptomen eine umfassende Diagnostik zu veranlassen, denn das Risiko einer falsch positiven Diagnose wird dadurch gleichzeitig erhöht. Dies hätte nicht nur eine Kostensteigerung, sondern auch eine erhöhtes Gefährdungspotential für die Patienten durch die nachfolgende, zum Teil invasive Diagnostik zur Folge.

Als Beispiel für eine schrittweise Diagnostik sei ein Patient mit persistierendem Husten seit zwei Wochen angeführt. Hier würde man die Lunge auskultieren und gegebenenfalls eine Messung des peak expiratory flow durchführen. Erst bei Beschwerdepersistenz lohnt sich eine Spirometrie und die Röntgendarstellung des Thorax. Eventuell kann nun ein Asthma bronchiale oder auch ein Tumor erkannt werden. Erst am Schluss stehen Bopyplethysmographie, Bronchoprovokation und Computertomographie zur Diagnostik bzw. Ausschluss von Asthma bronchiale, eines Lungenemphysems oder anderer Lungenerkrankungen. Dieses abgestufte Vorgehen lässt sich anhand eines evidenzbasierten diagnostischen Algorithmus abbilden, der eine effiziente hausärztliche Herangehensweise ermöglicht [21].

Problematisch bei der Einschätzung der diagnostischen Güte in der Primärversorgung sind Verzerrungen (engl: bias) und Variationen von Tests, die durch Selektion, Prävalenz und unterschiedlicher Schwere von Erkrankungen verursacht werden [14, 22]. Insbesondere die komplexe Interaktion verschiedener diagnostischer Faktoren und deren Bedeutung in der Allgemeinmedizin wurde bislang nur in wenigen Studien untersucht. Daher ist über die Auswirkungen der besonderen Arbeitsweise des Hausarztes mit unselektioniertem Patientenkontext noch wenig bekannt, so dass eine dringliche Notwendigkeit für diagnostische Studien in der Primärversorgung besteht [23].

Schlussfolgerungen

- Die Filterfunktion des Hausarztes mit der Notwendigkeit einer abgestuften Diagnostik lässt sich durch das Bayes'sche Theorem epidemiologisch gut erklären. Die Stärke der hausärztlichen Arbeitsweise liegt im Umgang mit Gesunden, die systemimmanente Schwäche liegt in der Diagnostik von Erkrankungen, die mit einer niedrigen Prävalenz vorkommen.
- Aufgrund der niedrigen PPV und hohen FPPV erscheint es problematisch, wenn im hausärztlichen Arbeitsfeld bei Vorliegen minimaler Symptome eine aufwendige Sofortdiagnostik gemacht wird, um möglichst viele Erkrankungen im Frühstadium zu erfassen. Der häufige Vorwurf, Hausärzte würden zu wenig diagnostizieren, kann vor diesem Hintergrund entkräftet werden. Hier wird der Arbeitsweise des Hausarztes nicht Rechnung getragen, die insbesondere dem bedeutsamen diagnostischen Faktor „Zeit“ gerecht wird.
- Die Integration von Zeit, klinischer Erfahrung und basaler apparativer Diagnostik ermöglicht eine effiziente Stufendiagnostik. Hierdurch ergibt sich eine Filterfunktion, wodurch die Beanspruchung der nachgeordneten Sektoren sinnvoll gesteuert wird.
- Die Schwächen der Diagnostik im Niedrigprävalenzbereich sind syste-

mimmanent. Das heißt, es gibt keine alternativen Gestaltungsmöglichkeiten für den diagnostischen Prozess. Denn die Effektivität im Hochprävalenzbereich der spezialärztlichen Versorgung lässt sich nur durch einen konsequenten „Filter“ primärärztlicher Versorgung sichern.

- Es besteht eine dringende Notwendigkeit zur Durchführung weiterer diagnostischer Studien, vor allem im Bereich der Primärversorgung, um die Versorgungsrealität und die Implikationen für das Gesundheitswesen adäquat abbilden zu können.

Literatur

- [1] Montnemery P, Hansson L, Lanke J, Lindholm LH, Nyberg P, Lofdahl CG et al. Accuracy of a first diagnosis of asthma in primary health care. *Fam Pract* 2002;19:365–368.
- [2] Steckelings UM, Stoppelhaar M, Sharma AM, Wittchen HU, Krause P, Kupper B et al. HYDRA: possible determinants of unsatisfactory hypertension control in German primary care patients. *Blood Press* 2004;13:80–88.
- [3] Kruse J, Schmitz N, Woller W, Heckrath C, Tress W. Warum übersieht der Hausarzt die psychischen Störungen seiner Patienten? *Psychother Psychosom Med Psychol* 2004;54:45–51.
- [4] Wittchen HU, Pittrow D. Prevalence, recognition and management of depression in primary care in Germany: the Depression 2000 study. *Hum Psychopharmacol* 2002;17 Suppl 1:S1–11.
- [5] Hensler S, Wiesemann A. Diskreditierende Versorgungsstudien in deutschen Hausarztpraxen. Oder: Der Versuch, die Prävalenz von Krankheiten und die medikamentöse Behandlungsbedürftigkeit zu steigern. *Z Allg Med* 2003;79:579–585.
- [6] O'Connor GT, Sox HC, Jr. Bayesian reasoning in medicine: the contributions of Lee B. Lusted, MD. *Med Decis Making* 1991;11:107–111.
- [7] Ledley RS, Lusted LB. Reasoning foundations of medical diagnosis; symbolic logic, probability, and value theory aid our understanding of how physicians reason. *Science* 1959;130:9–21.
- [8] Holleman DR, Jr., Simel DL, Goldberg JS. Diagnosis of obstructive airways disease from the clinical examination. *J Gen Intern Med* 1993;8:63–68.
- [9] Dodge RR, Burrows B. The prevalence and incidence of asthma and asthma-

like symptoms in a general population sample. *Am Rev Respir Dis* 1980;122:567–575.

- [10] Straus SE, McAlister FA, Sackett DL, Deeks JJ. The accuracy of patient history, wheezing, and laryngeal measurements in diagnosing obstructive airway disease. CARE-COAD1 Group. *Clinical Assessment of the Reliability of the Examination-Chronic Obstructive Airways Disease*. *JAMA* 2000;283:1853–1857.
- [11] Hunter CJ, Brightling CE, Woltmann G, Wardlaw AJ, Pavord ID. A comparison of the validity of different diagnostic tests in adults with asthma. *Chest* 2002;121:1051–1057.
- [12] Thiadens HA, de Bock GH, Dekker FW, Huysman JA, Van Houwelingen JC, Springer MP et al. Value of measuring diurnal peak flow variability in the recognition of asthma: a study in general practice. *Eur Respir J* 1998;12:842–847.
- [13] Remes ST, Pekkanen J, Remes K, Salonen RO, Korppi M. In search of childhood asthma: questionnaire, tests of bronchial hyperresponsiveness, and clinical evaluation. *Thorax* 2002;57:120–126.
- [14] Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Stat Med* 1997;16:981–991.
- [15] White KL, Williams TF, Greenberg BG. The ecology of medical care. *N Engl J Med* 1961;265:885–892.
- [16] Green LA, Fryer GE, Jr., Yawn BP, Lanier D, Dovey SM. The ecology of medical care revisited. *N Engl J Med* 2001;344:2021–2025.
- [17] Gerlach FM. Allgemeinmedizin in einem modernen Gesundheitssystem - Konsequenzen für Forschung, Lehre und Praxis (Teil 2). *Z ärztl Fortbild Qualsich* 2002;96:307–311.
- [18] Pewsner D, Battaglia M, Minder C, Marx A, Bucher HC, Egger M. Ruling a diagnosis in or out with "SpIn" and "SnNOut": a note of caution. *BMJ* 2004;329:209–213.
- [19] Müller-Bardorff M, Rauscher T, Kampmann M, Schoolmann S, Laufenberg F, Mangold D, Zerback R, Remppis A, Katus HA. Quantitative bedside assay for cardiac Troponin T: a complementary method to centralized laboratory testing. *Clinical Chemistry* 1999;45:1002–1008.
- [20] Abholz HH. Generalist und Spezialist - Zwei unterschiedliche Welten von Arbeit und Arbeitsauftrag. *Z Allg Med* 2004;80:411–413.
- [21] Schneider A, Borst MM, Gerlach FM, Szecsenyi J. Verdacht auf Asthma bron-

chiale – Entwicklung eines Algorithmus zur Stufendiagnostik in der Hausarztpraxis. *Z ärztl Fortbild Qualsich* 2003;97:485–493.

- [22] Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med* 2004;140:189–202.
- [23] Knottnerus JA, van Weel C, Muris JW. Evaluation of diagnostic procedures. *BMJ* 2002;324:477–480.

Korrespondenzadresse:

Dr. med. Antonius Schneider
Abteilung Allgemeinmedizin und
Versorgungsforschung
Universitätsklinikum Heidelberg
Voßstrasse 2, Gebäude 37
69115 Heidelberg
Tel.: 06221 / 564743
Fax: 06221 / 561972
E-Mail:
antonius.schneider@med.uni-heidelberg.de

