

Zwischen Evidenz und Praxis: wie bewerten Hausärzte Früherkennungsmaßnahmen? Eine Fragebogenstudie

Luca Frank, Susann Hueber, Thomas Kühlein, Angela Schedlbauer, Marco Roos

Angaben zur Veröffentlichung / Publication details:

Frank, Luca, Susann Hueber, Thomas Kühlein, Angela Schedlbauer, and Marco Roos. 2018. "Zwischen Evidenz und Praxis: wie bewerten Hausärzte Früherkennungsmaßnahmen? Eine Fragebogenstudie." Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen 135-136: 1-9.
<https://doi.org/10.1016/j.zefq.2018.06.003>.

Zwischen Evidenz und Praxis: Wie bewerten Hausärzte Früherkennungsmaßnahmen? Eine Fragebogenstudie

Evidence and practice: How do primary care physicians evaluate screening interventions? A questionnaire study

Luca Frank*, Susann Hueber, Thomas Kühlein, Angela Schedlbauer, Marco Roos

Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Allgemeinmedizinisches Institut, Erlangen, Germany

Z U S A M M E N F A S S U N G

Hintergrund: PatientInnen müssen vor Früherkennungsmaßnahmen über deren Nutzen und Risiken aufgeklärt werden. Diese Aufklärung stützt sich auf vorhandene Evidenz und verlangt von Ärztinnen und Ärzten Kompetenz im Umgang mit epidemiologischen Zahlen.

Fragestellung: Geben Hausärztinnen und Hausärzte anhand epidemiologischer Zahlen andere Empfehlungen zu Früherkennungsmaßnahmen, als sie es üblicherweise im Praxisalltag tun? Wie weit folgen diese Einschätzungen der Zahlen und das Vorgehen im Praxisalltag offiziellen Empfehlungen? Verändert sich die Einschätzung der Zahlen, wenn diese als Tabelle versus Piktogramm präsentiert werden?

Methoden: Onlinebefragung bayerischer Hausärztinnen und Hausärzte. Den Befragten wurden drei unterschiedliche Früherkennungs-Szenarien zur Einschätzung vorgelegt. Die Szenarien enthielten ausschließlich epidemiologische Zahlen aus Studien zu existierenden Früherkennungsmaßnahmen. Um vorgefasste Meinungen zu umgehen, wurden weder die Erkrankungen noch die Maßnahmen namentlich erwähnt. Die Darstellung erfolgte als Tabelle oder als Piktogramm. Anschließend wurde das Vorgehen zu denselben Früherkennungsmaßnahmen im Praxisalltag erfragt.

Ergebnisse: Antworten von 43 Hausärztinnen und Hausärzten wurden ausgewertet. Die Befragten würden die Früherkennungsmaßnahmen anhand der Zahlen in den Szenarien seltener empfehlen als im Praxisalltag ($F(1, 3) = 104,83; p = 0,002; \eta^2_{\text{partial}} = 0,97$). Im Praxisalltag folgten sie dabei überwiegend offiziellen Empfehlungen, während sie in den Szenarien eher von diesen abwichen. (Praxisalltag: $M = 1,79; SD = 0,77$; Szenarien: $M = 1,44; SD = 0,80; t(42) = -2,29; p = 0,03$). Die Darstellung hatte keinen Einfluss ($t(127) = -1,83; p = 0,07$).

Schlussfolgerung: Im Praxisalltag scheinen Zahlen aus der Evidenz zu Früherkennungsmaßnahmen nur wenig Bedeutung in der Entscheidungsfindung zu haben. Vermutlich würden sich die Hausärztinnen und Hausärzte bei Kenntnis der Zahlen häufiger gegen Früherkennungsmaßnahmen aussprechen.

A B S T R A C T

Background: Patients receiving a screening intervention have to be informed about risks and benefits. On the part of the physician, it requires the understanding of statistical evidence and statistical literacy.

Objectives: Do general practitioners (GPs) make different recommendations on screening interventions if they only rely on statistics compared to their decisions in everyday practice? Are the decisions relying on statistics and the decisions made in everyday practice consistent with official recommendations? Does the way of presenting the numbers (table versus pictogram) affect the decision?

* Korrespondenzadresse: Luca Frank, Universitätsklinikum Erlangen, Allgemeinmedizinisches Institut, Universitätsstr. 29, 91054 Erlangen, Germany. E-mail: luca.frank@uk-erlangen.de (L. Frank).

Methods: Online survey among German GPs. The GPs were asked to make recommendations for three different screening scenarios which were based on statistical evidence of existing screening interventions. To avoid bias, statistics were not presented in relation to the actual diseases. The numbers were presented in either a table or a pictogram. Afterwards, the GPs were asked for their recommendations on the same screening interventions in everyday practice.

Results: Forty-three GPs were surveyed. Compared to everyday medical practice, participants were less likely to recommend a screening intervention when being confronted with the underlying statistical evidence ($F(1, 3) = 104.83, p = .002, \eta^2_{\text{partial}} = .97$). Most of the decisions in everyday practice were consistent with official recommendations, while their decisions relying on statistics were more likely to deviate from them (everyday practice, $M = 1.79, SD = 0.77$; scenario, $M = 1.44, SD = 0.80$; $t(42) = -2.29, p = .03$). The way the numbers were presented did not affect the decision ($t(127) = -1.83, p = .07$).

Conclusions: In everyday practice, GPs' screening recommendations do not seem to be based on statistical evidence. Presumably they would be more reluctant to recommend screening interventions if they knew the statistical evidence.

Einleitung

Maßnahmen zur Früherkennung von Krankheiten und zur Gesundheitsförderung sind Grundelemente der medizinischen Versorgung gemäß §73 SGB V und des ärztlichen Handelns entsprechend §1 der Musterberufsordnung der Bundesärztekammer [1]. Die Beratung zur Früherkennung von Krankheiten gehört zu den Aufgaben der Hausärztinnen und Hausärzte. Oft fordern PatientInnen von ihnen, persönlich eine Empfehlung für oder gegen eine Früherkennungsmaßnahme auszusprechen. Hierfür müssen Ärztinnen und Ärzte die Ergebnisse von Studien bewerten und die Erkenntnisse in den Praxisalltag übertragen. Das Deutsche Netzwerk Evidenzbasierte Medizin (DNEbM) definiert dies als „Integration individueller klinischer Expertise mit der bestmöglichen externen Evidenz aus systematischer Forschung (modifiziert nach [2])“ im Sinne einer evidenzbasierten Medizin (EbM) [3]. In einer qualitativen Erhebung unter deutschen Hausärztinnen und Hausärzten zeigte sich allerdings, dass diese Schwierigkeiten bei der Umsetzung von EbM im Alltag haben, obwohl sie ihr eigentlich positiv gegenüberstehen [4]. Besonders der hohe Zeitaufwand bei der Recherche und Bewertung von Evidenz ist dabei eine Barriere. In einer weiteren Studie an deutschen Hausärztinnen und Hausärzten sahen einige gar ihre persönliche Berufserfahrung, im Sinne einer „internen Evidenz“, als ausreichend zur Beratung von PatientInnen an [5]. In diesem Kontext stellt sich die Frage, ob Hausärztinnen und Hausärzte im Alltag eigentlich die Evidenz heranziehen, um eine Empfehlung für oder gegen eine Früherkennungsmaßnahme auszusprechen?

Die Entscheidung für oder gegen eine Früherkennungsmaßnahme ist eine komplexe Aufgabe. Leitlinien und systematische Übersichtsarbeiten von Fachgesellschaften und wissenschaftlichen Organisationen können hier Hilfestellung geben. Auf Basis des in Studien gefundenen Nutzens und Schadens von Früherkennungsmaßnahmen werden offizielle Empfehlungen für oder gegen eine bestimmte Methode ausgesprochen. So gibt es eine offizielle Empfehlung zur Durchführung der Endoskopie zur Darmkrebsfrüherkennung. Sie stellt damit ein Verfahren mit anerkanntem Nutzen dar [6,7]. Das PSA-Screening hingegen gilt als ein Verfahren mit ungünstigem Nutzen-Schaden-Verhältnis [8,9], weshalb sich die Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM) gegen ein PSA-Screening zur Früherkennung von Prostatakrebs ausspricht [10,11]. Bei der Mammografie herrscht Uneinigkeit: Sie gilt als Früherkennungsmaßnahme mit noch nicht eindeutiger Datenlage [12–14]. Während das Mammografiescreening in Deutschland als nationales Früherkennungsprogramm durchgeführt wird, rät das Swiss Medical Board von solchen Programmen ab [15]. Über die genannten Hilfestellungen hinaus benötigen Ärztinnen und Ärzte aber selbst ein ausreichendes Verständnis der zugrunde liegenden Zahlen aus Studien. Dies ist bei der Aufklärung von PatientInnen im Sinne einer informierten

Einwilligung (informed consent) von besonderer Bedeutung. Nur bei eigener Kenntnis der Evidenz ist eine individualisierte Beratung möglich. Auch Hausärztinnen und Hausärzte betonten in einer qualitativen Studie aus Deutschland die Bedeutung einer individuellen Prävention [16]. Ob aktuelle Leitlinien diesem Anspruch gerecht werden, ist fraglich. In einer Erhebung unter Hausärztinnen und Hausärzten warfen diese Leitlinien eine mangelnde Hausarztzentrierung und Übertragbarkeit auf die Praxis vor [4]. Dies wirft die Frage auf, wieweit Hausärztinnen und Hausärzte im Alltag offiziellen Empfehlungen zu Früherkennungsmaßnahmen folgen und welche Rolle ihre eigene Einschätzung der Evidenz spielt?

Um zu einer gemeinsamen Entscheidung zu gelangen, müssen PatientInnen und Ärztinnen und Ärzte Nutzen und Risiken einer Früherkennungsmaßnahme kennen, verstehen und gegeneinander abwägen können. Dabei geben epidemiologische Zahlen Informationen zu Nutzen und Risiken von Früherkennungsmaßnahmen. Die Fähigkeit zur Interpretation epidemiologischer Zahlen wie Mortalität, Morbidität und Number-Needed-to-Screen (NNS) wird als statistische Kompetenz (englisch: statistical literacy) bezeichnet. Diese Kompetenz haben Ärztinnen und Ärzte selten in der Ausbildung vermittelt bekommen. Auch heute ist sie kaum Gegenstand von Lehre und Prüfung. Im Rahmen einer repräsentativen Befragung von bayerischen Absolventen der Humanmedizin gaben nur 30,4% der Befragten an, nötige Kompetenz zum wissenschaftlichen Handeln, wie beispielsweise zur Bewertung von Studien, im Studium erworben zu haben [17]. Deshalb ist es kein Wunder, dass Studien zur statistischen Kompetenz von Ärztinnen und Ärzten auf Defizite in diesem Bereich hinweisen. So wurden beispielsweise deutsche GynäkologInnen zu einer typischen Problemstellung aus ihrem Berufsalltag befragt [18]. Sie sollten den positiven prädiktiven Wert einer Mammografie aus epidemiologischen Zahlen errechnen. Die Mehrheit der Befragten scheiterte an dieser Aufgabe. Eine Befragung von US-amerikanischen Primärmedizinerinnen und -medizinern anhand hypothetischer Früherkennungsszenarien ergab zudem, dass es ihnen schwerfällt, relevante von irrelevanten Zahlen zu unterscheiden [19].

Ein wichtiger Aspekt beim Umgang mit Zahlen aus der Evidenz ist darüber hinaus die Darstellung dieser Zahlen. So kann eine transparente Darstellung der Zahlen diese auch Lesern mit geringer statistischer Kompetenz verständlich nahe bringen [20]. Eine aktuelle Leitlinie des DNEbM setzt hier Standards für die evidenzbasierte Gesundheitsinformation für PatientInnen [21]. In ihr wird betont, dass Nutzen und Schaden durch absolute Risikomaße dargestellt werden sollen [21]. Eine Darstellung in relativen Risikomaßen hingegen führt häufig zu einer Fehleinschätzung von Nutzen und Schaden, sowohl bei PatientInnen als auch Ärztinnen und Ärzten [18,22]. Für bestmögliche Verständlichkeit muss dabei jedoch auf die Verwendung gleicher Bezugsgrößen geachtet werden [21,23]. Zudem ist die Darstellung als absolute Risikoreduktion

(ARR) für PatientInnen leichter verständlich als eine Number-Needed-to-Treat (NNT) [21,24]. Dieser Unterschied verschwand aber bei zusätzlicher Angabe des entsprechenden Basisrisikos. Insbesondere für Ärztinnen und Ärzte hat die NNT trotzdem einen Mehrwert im Alltag [25]. Die Effizienz von Therapien lässt sich einfacher vergleichen und merken als eine ARR. Zusätzlich können gemäß der Leitlinie des DNEbM Grafiken, insbesondere Piktogramme, unterstützend bei der Darstellung von Risiken eingesetzt werden [21]. Diese Art der Darstellung findet sich beispielsweise auch in der Handlungsempfehlung zum PSA-Screening der DEGAM [10]. Eine Untersuchung an Studierenden konnte für die Darstellung in Piktogrammen eine Verbesserung des Verständnisses statistischer Zahlen nachweisen [26]. Im Rahmen einer qualitativen Untersuchung mit Ärztinnen und Ärzten wurden Piktogramme zum Thema PSA-Screening getestet [27]. Die Teilnehmenden gaben unter anderem an, dass sich durch die Piktogramme ihr Verständnis für die Effekte des PSA-Screenings verbessert hat. Darüber hinaus ist die Bedeutung solcher Entscheidungshilfen für Ärztinnen und Ärzte aber nach wie vor kaum erforscht. Fraglich ist, ob auch objektiv eine Veränderung der Entscheidungen der Ärztinnen und Ärzte messbar ist.

Mit der vorliegenden Studie untersuchten wir, ob Ärztinnen und Ärzte zu bestimmten Früherkennungsmaßnahmen anhand epidemiologischer Zahlen, also der Evidenz, andere Empfehlungen geben würden, als sie es üblicherweise in der Praxis tun. Die Fragestellungen waren:

- (1) Führt die alleinige Präsentation epidemiologischer Zahlen zu einer Abweichung der Einschätzung einer Früherkennungsmaßnahme vom bisherigen Vorgehen im Praxisalltag?
- (2) Wie weit folgen diese Einschätzung und das bisherige Vorgehen im Praxisalltag offiziellen Empfehlungen?
- (3) Hat die Art und Weise, wie epidemiologische Zahlen präsentiert werden (Tabelle versus Piktogramm), einen Einfluss auf die Einschätzung von Früherkennungsmaßnahmen?

Methoden

Design und Stichprobe

Aufgrund bereits bestehender qualitativer Studien zum Thema der Anwendung von Evidenz und Früherkennungsmaßnahmen in der Hausarztpraxis, sowie eingeschränkter Forschungsressourcen im Rahmen einer medizinischen Dissertation, wurde auf eine qualitative Vorarbeit verzichtet [4,5,16,27]. Es wurde eine quantitative Querschnittserhebung mittels eines Online-Fragebogens durchgeführt. Dieser wurde mit der Software SurveyMonkey erstellt und von zwei ärztlich tätigen Mitarbeitern (AD, AS) getestet. Die Unbedenklichkeitsbescheinigung der Ethik-Kommission der Medizinischen Fakultät der Friedrich-Alexander-Universität Erlangen-Nürnberg liegt seit November 2014 vor (327_14c).

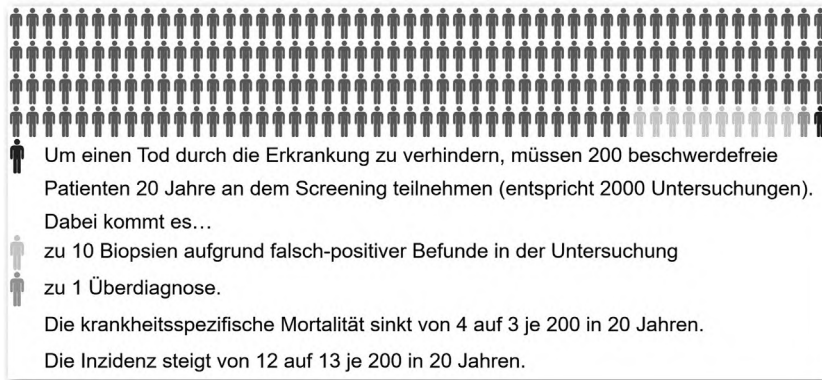
Hausärztinnen und Hausärzte der Lehrpraxen der Friedrich-Alexander-Universität (n=88) und der Jungen Allgemeinmedizin Deutschland (JADE, n=941) wurden Ende Juli 2015 per Email zur Online-Befragung eingeladen, ein Reminder wurde eine Woche später Anfang August versendet. Es gab keine Entschädigung für die Teilnahme an der Studie. Um Mehrfachteilnahmen auszuschließen, wurde eine IP-Sperre verwendet. Es nahmen 72 hausärztlich tätige Ärztinnen und Ärzte an der Umfrage teil (Rücklauf=7%). Von diesen konnten 43 vollständig beantwortete Fragebögen ausgewertet werden. Die Stichprobe bestand vor allem aus Fachärztinnen und -ärzten für Allgemeinmedizin (63%), daneben aus Ärztinnen und Ärzten in Weiterbildung zum Facharzt Allgemeinmedizin (30%) und einer kleinen Zahl an Fachärztinnen und -ärzten für Innere Medizin (7%). Eine Vielzahl der Teilnehmenden war erst seit weniger als fünf Jahren (49%) und überwiegend (63%) in Berufsausübungsgemeinschaften (Gemeinschaftspraxen, MVZs, usw.) tätig. Über die Hälfte der Befragten (63%) gab an, mehr als fünf Mal in einer Arbeitswoche ein Gespräch zu Früherkennungsmaßnahmen zu führen. Die vollständigen demografischen Angaben sind in [Tabelle 1](#) zusammengefasst.

Tabelle 1
Demografische Merkmale der Befragten (n=43).

Merkmal	Junge Hausärztinnen und Hausärzte Weniger als 5 Jahre hausärztlich tätig	Erfahrene Hausärztinnen und Hausärzte Mehr als 5 Jahre hausärztlich tätig	Gesamt	Hausärztinnen und Hausärzte Deutschland [33]
Anzahl (n)	21	22	43	
Alter (Jahre)	36* (± 4)	52* (± 10)	44 (± 11)	55
Geschlecht (%)				
Männlich	29*	77*	53	56
Weiblich	71*	23*	47	44
Berufsbezeichnung (%)				
Facharzt für Allgemeinmedizin	29*	95*	63	63
Facharzt für Innere Medizin	10*	5*	7	28
Arzt in Weiterbildung	62*		30	
Praktischer Arzt				9
Praxisform (%)				
Einzelpraxis	29	32	30	43
Praxis mit Kollegen	57	68	63	57
Im Krankenhaus	14		7	
Beratungshäufigkeit (%)				
Weniger als 5 Mal pro Woche	52*	23*	37	
Mehr als 5 Mal pro Woche	48*	77*	63	

Signifikante Unterschiede der Gruppen in χ^2 -Tests bzw. T-Tests wurden mit einem * gekennzeichnet. Im Vergleich werden Angaben zu Hausärztinnen und Hausärzten in Deutschland aus dem Jahre 2017 aufgeführt [33].

Piktogramm



Tabelle

Number-needed-to-screen:	Teilnahme von 200 Personen für 20 Jahre (entspricht 2000 Untersuchungen)
Falsch-Positiv-Rate:	10 Biopsien je 200 in 20 Jahren aufgrund falsch-positiver Befunde in der Untersuchung
Überdiagnosen:	1 je 200 in 20 Jahren
Krankheitsspezifische Mortalität:	Ungescreent: 4 je 200 in 20 Jahren Gescreent: 3 je 200 in 20 Jahren
Inzidenz:	Ungescreent: 12 je 200 in 20 Jahren Gescreent: 13 je 200 in 20 Jahren

Abbildung 1. Früherkennungsszenarien: Darstellung der epidemiologischen Zahlen als Piktogramm (oberer Teil) und Tabelle (unterer Teil). Abbildung erstellt aus Screenshot des verwendeten Fragebogens.

Messinstrument

Um die Einschätzung von Früherkennungsmaßnahmen allein anhand epidemiologischer Zahlen zu messen, wurden in Anlehnung an Wegwarth et al. drei Früherkennungsszenarien konzipiert [19]. Als Grundlage der Szenarien dienten die epidemiologischen Zahlen zu Früherkennungsmaßnahmen der drei häufigsten Krebserkrankungen in Deutschland: der Endoskopie zur Früherkennung von Darmkrebs, der Mammografie für Brustkrebs und dem PSA-Screening für Prostatakrebs [6,9,10,14,28]. Um sicherzustellen, dass die Entscheidung der Befragten ausschließlich auf den Zahlen beruht, wurden weder die Erkrankungen noch die Untersuchungsmethoden in den Szenarien namentlich erwähnt. Jedes Szenario begann mit folgendem Einleitungstext: „Ein 58-jähriger gesunder Patient ohne relevante Risikofaktoren kommt in Ihre Sprechstunde und bittet Sie um Rat bezüglich einer spezifischen Untersuchung zur Früherkennung einer Tumorerkrankung. Der Patient vertraut auf Ihre Meinung und möchte wissen, ob Sie ihm diese Untersuchung empfehlen können. Sie recherchieren zunächst, welche Erkenntnisse es zu Nutzen und Risiken gibt, und finden folgende Informationen aus verlässlichen wissenschaftlichen Studien.“ Es folgten die Zahlen zu Inzidenz und Mortalität der Erkrankung in Screening- und Kontroll-Gruppe, sowie zu Number-Needed-to-Screen (NNS) und Schadenspotential (Falsch-Positiv-Rate, Komplikationen, Überdiagnosen). Die Zahlen wurden ausgewählt, da sie typischerweise in Studien zu Früherkennungsmaßnahmen berichtet werden und alle außer der Inzidenz eine hohe Relevanz bei der Bewertung haben. Die Darstellung der Zahlen erfolgte in natürlichen Häufigkeiten. Prozentuale Risiken in der Literatur wurden gegebenenfalls in diese überführt. Es wurde auf möglichst einheitliche Bezugsgrößen innerhalb der einzelnen

Szenarien geachtet. Die Number-Needed-to-Screen wurde, wo nicht bereits in der Literatur angegeben, aus der absoluten Risikoreduktion der krankheitsspezifischen Mortalität berechnet. Anhand der Zahlen wurden zusätzlich mit Photoshop Piktogramme in Anlehnung an die Handlungsempfehlung zum PSA-Screening der DEGAM erstellt [10]. Die Piktogramme und Tabellen enthielten stets dieselben Zahlen mit identischen Bezugsgrößen. Die Formulierungen in beiden Darstellungen waren nur geringfügig abweichend. Beide Darstellungsformen sind in [Abbildung 1](#) zu sehen.

Durchführung

Der Fragebogen untergliederte sich in zwei Teile. Im ersten Teil wurde mittels der Szenarien die Einschätzung von Früherkennungsmaßnahmen allein anhand epidemiologischer Zahlen erhoben. Jeder Befragte musste alle drei Szenarien einzeln und in zufälliger Reihenfolge einmal bewerten. Nach jedem Szenario waren die Befragten aufgefordert, eine Empfehlung auf einer 6-stufigen Likert-Skala abzugeben (Ich empfehle die Untersuchung: 1 = „auf jeden Fall“ bis 6 = „auf gar keinen Fall“). Im Fragebogen konnte nicht zum vorherigen Szenario zurückgesprungen werden, um eine nachträgliche Korrektur der Antworten auszuschließen. Die epidemiologischen Zahlen wurden dabei für jedes Szenario zufällig entweder als Piktogramm oder Zahlentabelle dargeboten. Ein Befragter konnte also beispielsweise zwei Szenarien als Piktogramm und eines als Tabelle präsentiert bekommen. Die Randomisierung der Befragten in den drei Szenarien führte zu annähernd gleich großen Gruppen mit ähnlicher demographischer Zusammensetzung. [Tabelle 2](#) gibt eine Übersicht über die Verteilung der Merkmale in den Gruppen.

Tabelle 2
Demografische Merkmale der Befragten (n = 43) in den Darstellungsgruppen der einzelnen Szenarien.

Merkmal	Szenario 1		Szenario 2		Szenario 3		Alle Szenarien	
	Tabelle	Piktogramm	Tabelle	Piktogramm	Tabelle	Piktogramm	Tabelle	Piktogramm
Anzahl (n)	25	18	22	21	24	19	71	58
Alter (Jahre)	46 (± 11)	41 (± 10)	43 (± 10)	45 (± 11)	43 (± 11)	45 (± 11)	44 (± 11)	44 (± 11)
Geschlecht (%)								
Männlich	64	39	50	57	42	68	52	55
Weiblich	36	61	50	43	58	32	48	45
Berufsbezeichnung (%)								
Facharzt für Allgemeinmedizin	72	50	55	71	54	74	61	66
Facharzt für Innere Medizin	8	6	5	10	4	11	6	9
Arzt in Weiterbildung	20	44	41	19	42	16	34	26
Praktischer Arzt								
In hausärztlicher Tätigkeit (%)								
Weniger als 5 Jahre	44	56	55	43	50	47	49	48
Mehr als 5 Jahre	56	44	45	57	50	53	51	52
Praxisform (%)								
Einzelpraxis	28	33	27	33	33	26	30	31
Praxis mit Kollegen	68	56	68	57	58	68	65	60
Im Krankenhaus	4	11	5	10	8	5	6	9
Beratungshäufigkeit (%)								
Weniger als 5 Mal pro Woche	28	50	36	38	58*	11*	41	33
Mehr als 5 Mal pro Woche	72	50	64	62	42*	89*	59	67

Signifikante Unterschiede der Gruppen in χ^2 -Tests bzw. T-Tests wurden mit einem * gekennzeichnet.

Anschließend sollte im zweiten Teil des Fragebogens das tatsächliche, im Praxisalltag stattfindende Empfehlungsverhalten eruiert werden. Den Befragten wurde dafür die Frage gestellt: „Wenn Sie an Ihren Praxisalltag denken: Welche Untersuchungen würden Sie einem 58-jährigen gesunden Patienten ohne relevante Risikofaktoren empfehlen?“ Darunter sollten sie ebenfalls auf einer 6-stufigen Likert-Skala angeben, ob sie die Früherkennungsmaßnahmen Mammografie, PSA-Test, Endoskopie zur Darmkrebsfrüherkennung, Fäkaler okkultur Bluttest, PAP-Abstrich und Hautinspektion empfehlen würden. Die Untersuchungsverfahren und entsprechenden Erkrankungen wurden im Gegensatz zu den Szenarien jetzt explizit benannt, ohne jedoch epidemiologische Zahlen aufzuführen. Die drei zusätzlich erfragten Früherkennungsmaßnahmen sollten verhindern, dass die Anzahl der erfragten Maßnahmen mit der Anzahl der Szenarien übereinstimmt und so zu einer Anpassung der Antworten führt. Sie waren nicht Teil der Auswertung.

Datenanalyse

Die Auswertung erfolgte mit SPSS Statistics Version 23. Zum Vergleich der Einschätzungen in den Szenarien mit dem Vorgehen im Praxisalltag wurde eine zweifaktorielle Varianzanalyse mit Messwiederholung auf den beiden Innersubjektfaktoren „Art der Einschätzung“ (im Szenario, im Praxisalltag) und „Früherkennungsmaßnahme“ (Darmkrebsfrüherkennung, Mammografie, PSA-Screening) durchgeführt. Dabei wurden die Zwischensubjektfaktoren Alter, Geschlecht, Berufsbezeichnung, Erfahrung in hausärztlicher Tätigkeit, Praxisform und Beratungshäufigkeit mit berücksichtigt. Bei signifikanten Effekten erfolgten zudem Einzelvergleiche mittels T-Tests. Zusätzlich wurden die beiden Gruppen junge (weniger als fünf Jahre hausärztlich tätig) und erfahrene Hausärztinnen und Hausärzte (mehr als fünf Jahre hausärztlich tätig) bezüglich ihrer mittleren Einschätzung sowohl in den Szenarien als auch beim Vorgehen im Praxisalltag mittels T-Tests explorativ untersucht.

Dann wurden die Einschätzungen zu den Szenarien und zum Vorgehen im Praxisalltag mit offiziellen Empfehlungen verglichen. Die verwendeten Früherkennungsmaßnahmen (PSA-Test, Mammografie und Endoskopie zur Darmkrebsfrüherkennung) unterscheiden sich deutlich hinsichtlich ihrer offiziellen Empfehlung. Wie bereits erwähnt, gilt die Darmkrebsfrüherkennung als ein Verfahren mit anerkanntem Nutzen [6,7], die Mammografie

als ein Verfahren mit noch nicht eindeutiger Datenlage [12–15] und das PSA-Screening als ein Verfahren mit ungünstigem Nutzen-Schaden-Verhältnis [8–11]. Die Antworten auf der 6-stufigen Likert-Skala wurden bei Werten am unteren Ende der Skala (1–2) als klar positive Einschätzung einer Früherkennungsmaßnahme, bei Werten am oberen Ende der Skala (5–6) als klar negative Einschätzung und bei Werten in der Mitte der Skala (3–4) als nicht eindeutig positive oder negative Einschätzung eingeordnet. Bei den verschiedenen Früherkennungsmaßnahmen wurden dementsprechend folgende Einschätzungen als der offiziellen Empfehlung folgend kodiert:

- | | |
|----------------------------|--|
| (1) Darmkrebsfrüherkennung | Werte 1–2: Untersuchung wird empfohlen |
| (2) Mammografie | Werte 3–4: Keine eindeutige Empfehlung |
| (3) PSA-Screening | Werte 5–6: Untersuchung wird nicht empfohlen |

Es wurde die Anzahl der Antworten, die offiziellen Empfehlungen folgten, jeweils in den Szenarien und im Praxisalltag, für jeden Befragten aufsummiert und separat mittels T-Test verglichen. Für die einzelnen Früherkennungsmaßnahmen erfolgte ferner ein Vergleich der Antworten, die offiziellen Empfehlungen folgten, zwischen Szenarien und Praxisalltag mittels McNemar-Tests.

Um herauszufinden, ob sich die Einschätzungen über alle drei Szenarien hinweg in Abhängigkeit von der Darstellungsweise unterscheiden, wurde ein T-Test für unverbundene Stichproben verwendet. Der Einfluss der Darstellungsweise auf die Einschätzungen in den einzelnen Szenarien wurde ebenfalls mittels T-Tests untersucht.

Ergebnisse

Vergleich Früherkennungsszenarien versus Vorgehen im Praxisalltag

Die zweifaktorielle Varianzanalyse zeigte, dass die Befragten die Früherkennungsmaßnahmen in den Szenarien signifikant seltener empfehlen würden als im Praxisalltag (Szenarien: $M=4,11$; $SD=1,34$; Praxisalltag: $M=3,29$; $SD=1,72$; $F(1, 3)=104,83$; $p=0,002$; $\eta^2_{\text{partial}}=0,97$). Die Unterschiede der verschiedenen Früherkennungsmaßnahmen untereinander waren ebenfalls signifikant ($F(2, 6)=112,36$; $p<0,001$; $\eta^2_{\text{partial}}=0,97$). Dabei gab es eine signifikante Wechselwirkung „Art der Einschätzung“

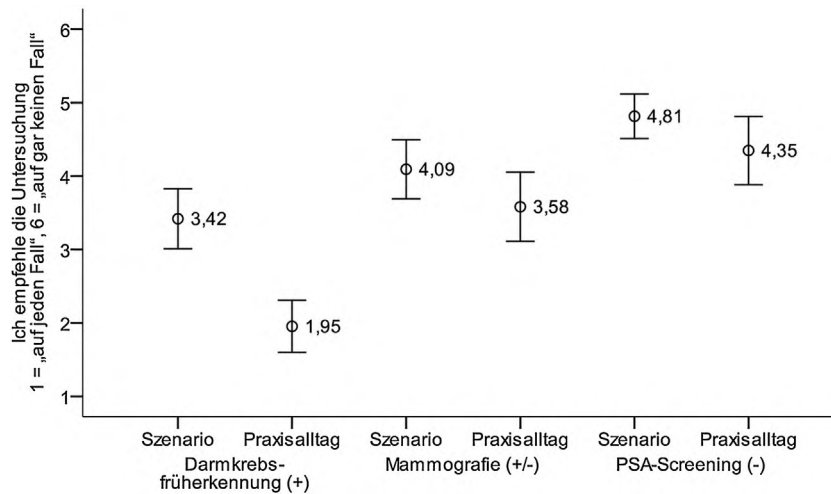


Abbildung 2. Verteilung der Einschätzungen im Szenario und beim Vorgehen im Praxisalltag. Es sind die Mittelwerte mit den entsprechenden 95%-Konfidenzintervallen dargestellt. Die Früherkennungsmaßnahmen wurden entsprechend ihrer positiven (+), uneindeutigen (+/-) oder negativen (-) Nutzen-Schaden-Bewertung in offiziellen Empfehlungen markiert.

X „Früherkennungsmaßnahme“ ($F(2, 6) = 8,31; p = 0,02; \eta^2_{\text{partial}} = 0,74$). Der Unterschied zwischen der Einschätzung im Szenario und dem Vorgehen im Praxisalltag war bei der Darmkrebsfrüherkennung stets größer als bei den anderen beiden Früherkennungsmaßnahmen. Die Wechselwirkung war ordinal und beide Haupteffekte blieben somit sinnvoll interpretierbar. Zudem zeigte sich ein signifikanter Einfluss der Beratungshäufigkeit ($F(1, 3) = 25,82; p = 0,01; \eta^2_{\text{partial}} = 0,90$). Befragte, die weniger als fünf Mal pro Woche zu Früherkennungsmaßnahmen beraten, würden diese sowohl in den Szenarien als auch im Praxisalltag seltener empfehlen.

Die Einzelvergleiche der Früherkennungsmaßnahmen bestätigten die Ergebnisse der Varianzanalyse. Der Großteil der Befragten gab an, die Darmkrebsfrüherkennung im Praxisalltag zu empfehlen, wohingegen ihre Einschätzungen der Zahlen im Szenario eher nicht eindeutig ausfielen (Szenario: $M = 3,42; SD = 1,33$; Praxisalltag: $M = 1,95; SD = 1,15$; $t(42) = 5,87; p < 0,001$). Auch das PSA-Screening würden die Befragten im Szenario tendenziell weniger empfehlen als im Praxisalltag (Szenario: $M = 4,81; SD = 0,98$; Praxisalltag: $M = 4,35; SD = 1,51$; $t(42) = 2,15; p = 0,04$). Gleiches galt für das Mammografiescreening, wobei der Unterschied hier nicht signifikant war (Szenario: $M = 4,09; SD = 1,31$; Praxisalltag: $M = 3,58; SD = 1,53$; $t(42) = 1,53; p = 0,13$). Eine grafische Übersicht der Einschätzungen gibt [Abbildung 2](#).

Zwischen jungen (weniger als fünf Jahre hausärztlich tätig) und erfahrenen Hausärztinnen und Hausärzten (mehr als fünf Jahre hausärztlich tätig) zeigte sich kein Unterschied in den Einschätzungen in den Szenarien (Junge: $M = 4,02; SD = 0,78$; Erfahrene: $M = 4,20; SD = 0,95$; $t(41) = -0,68; p = 0,50$). Gleiches galt für das Vorgehen im Praxisalltag (Junge: $M = 3,22; SD = 0,93$; Erfahrene: $M = 3,36; SD = 0,90$; $t(41) = -0,51; p = 0,62$).

Vergleich mit offiziellen Empfehlungen - Früherkennungsszenarien versus Vorgehen im Praxisalltag

Insgesamt wick mehr als die Hälfte der Einschätzungen anhand der Zahlen in den Szenarien von den entsprechenden offiziellen Empfehlungen ab (52%). Das Vorgehen im Praxisalltag folgte signifikant häufiger offiziellen Empfehlungen als die Einschätzungen in den Szenarien (Praxisalltag: $M = 1,79; SD = 0,77$; Szenarien: $M = 1,44; SD = 0,80$; $t(42) = -2,29; p = 0,03$). Bei der Darmkrebsfrüherkennung folgten die Befragten im Praxisalltag überwiegend offiziellen Empfehlungen, während die Mehrzahl der Einschätzungen

im Szenario von diesen abwich (Szenario = 33%; Praxisalltag = 77%; $p < 0,01$). Beim Mammografiescreening folgte sowohl im Szenario als auch im Praxisalltag etwa die Hälfte der Befragten offiziellen Empfehlungen (beide 44%; $p = 1,00$). Beim PSA-Screening folgten die Befragten im Szenario als auch im Praxisalltag überwiegend der kritischen Haltung der offiziellen Empfehlungen, wobei es im Szenario etwas mehr waren (Szenario = 67%; Praxisalltag = 58%; $p = 0,45$). [Abbildung 3](#) verdeutlicht die Ergebnisse grafisch.

Einfluss der Darstellung

In den drei Szenarien führte die Darstellung als Piktogramm ($n = 129$) dazu, dass die Früherkennungsmaßnahmen im Vergleich zur Darstellung als Zahlentabelle tendenziell seltener empfohlen wurden. Der Unterschied war jedoch nicht signifikant (Piktogramm: $M = 4,34; SD = 1,33$; Zahlentabelle: $M = 3,92; SD = 1,32$; $t(127) = -1,83; p = 0,07$). Die Ergebnisse bestätigten sich in den Einzelvergleichen. Bei der Darmkrebsfrüherkennung empfahlen die Befragten die Früherkennungsmaßnahme tendenziell seltener anhand des Piktogramms, ohne dass dieser Unterschied signifikant war (Piktogramm: $M = 3,72; SD = 1,36$; Zahlentabelle: $M = 3,20; SD = 1,29$; $t(41) = -1,28; p = 0,21$). Beim Mammografiescreening verhielten sich die Einschätzungen ebenso (Piktogramm: $M = 4,33; SD = 1,39$; Zahlentabelle: $M = 3,86; SD = 1,21$; $t(41) = -1,19; p = 0,24$). Gleiches galt für das PSA-Screening (Piktogramm: $M = 4,95; SD = 0,97$; Zahlentabelle: $M = 4,71; SD = 1,00$; $t(41) = -0,79; p = 0,44$).

Diskussion

Die Befragten empfahlen die Früherkennungsmaßnahmen anhand der epidemiologischen Zahlen in den Szenarien seltener, als sie dies im Praxisalltag tun. Im Praxisalltag folgten sie eher den offiziellen Empfehlungen. Die Art der Darstellung der Zahlen, als Piktogramm im Vergleich zur Zahlentabelle, hatte keinen Einfluss auf die Bewertung.

Legt man den befragten Hausärztinnen und Hausärzten epidemiologische Zahlen zur Bewertung der Sinnhaftigkeit von Früherkennungsmaßnahmen vor, so widerspricht ihre Schlussfolgerung häufig ihrem bisherigen Rat den PatientInnen gegenüber im Praxisalltag. Sie würden die Früherkennungsmaßnahmen anhand der Zahlen deutlich seltener empfehlen. Die Entscheidung, eine Früherkennungsmaßnahme zu empfehlen oder nicht, treffen die

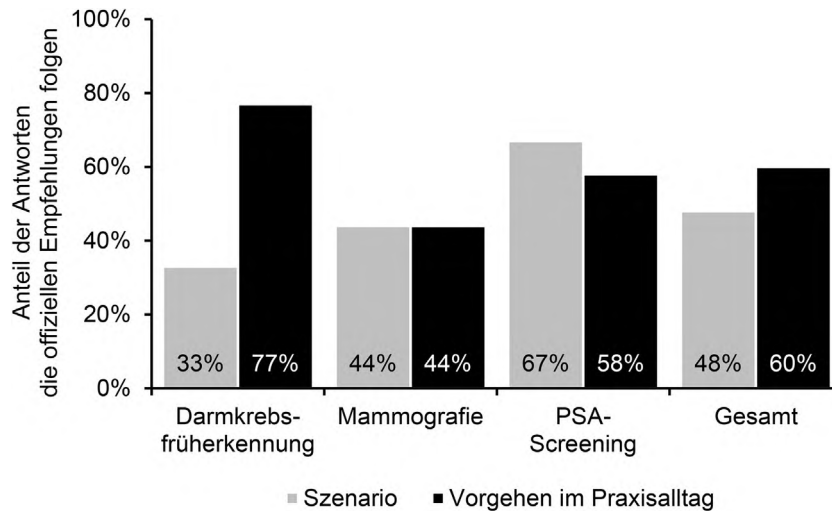


Abbildung 3. Anteil der Antworten in der Stichprobe (n=43), die offiziellen Empfehlungen folgen, in den Szenarien (graue Balken) und beim Vorgehen im Praxisalltag (schwarze Balken).

Befragten im Praxisalltag vermutlich nicht auf Basis der Zahlen aus Studien. Grund hierfür könnte sein, dass Hausärztinnen und Hausärzte häufig nicht mit aktuellen Studienergebnissen vertraut sind [29]. Dies würde die Ergebnisse vorangegangener qualitativer Untersuchungen bestätigen. Obwohl Hausärztinnen und Hausärzte sich allgemein dem Thema EbM verpflichtet fühlen, fällt ihnen die Umsetzung in der Praxis schwer [4]. Empfehlungen zu Früherkennungsmaßnahmen gründen dementsprechend selten auf der Evidenz [16]. Problematisch ist zudem, dass sich dieses Muster scheinbar auch schon bei jungen Hausärztinnen und Hausärzten findet. Bei den jungen Befragten mit weniger als fünf Jahren hausärztlicher Erfahrung zeigten sich dieselben Befunde, wie bei den erfahreneren KollegInnen. Als Grund für die seltene Anwendung der EbM in der Praxis wird in anderen Studien angeführt, dass es im stressigen Praxisalltag häufig schwer fällt, nach aktueller Evidenz zu suchen [4,30]. So kommen neue Studienergebnisse nur langsam in der Praxis an. Der Zugang zu verlässlichen, aktuellen Zahlen sollte Ärztinnen und Ärzten deshalb erleichtert werden. Hier eröffnen sich neue Möglichkeiten durch die zunehmende Verbreitung von digitalen Entscheidungshilfen für PatientInnen. Entscheidungshilfen wie der kardiovaskuläre Risikorechner arriba oder die Handlungsempfehlung zum PSA-Screening der DEGAM können nicht nur PatientInnen, sondern nebenbei auch Ärztinnen und Ärzte mit aktuellen Zahlen aus Studien versorgen [10,31].

Die Darstellung der Zahlen als Piktogramm im Vergleich zur Zahlentabelle hatte keinen Einfluss auf die Einschätzungen in den Szenarien. Das mag daran liegen, dass es durch die Darstellung als Piktogramm zu keiner Verbesserung des Verständnisses der epidemiologischen Zahlen kommen konnte. Dies bestätigt Ergebnisse aus anderen Studien an PatientInnen, die in der Leitlinie evidenzbasierte Gesundheitsinformation zusammengefasst wurden [21]. Für ergänzende grafische Darstellungen, wie Piktogramme, konnte insgesamt kein relevanter Effekt auf kognitive Endpunkte im Vergleich zu alleinigen Darstellungen der Zahlen in Tabellen nachgewiesen werden. Das scheint auch für Ärztinnen und Ärzte zu gelten. Ergebnisse einer vorangegangenen qualitativen Untersuchung an Ärztinnen und Ärzten, in der das selbsteingeschätzte Verständnis durch Piktogramme verbessert wurde, konnten somit nicht bestätigt werden [27]. In einer anderen Studie konnte zwar gezeigt werden, dass grafische Darstellungen in Form von Piktogrammen das Verständnis statistischer Zahlen bei Laien verbessern [26]. Allerdings reduzierte sich der Effekt deutlich, wenn die Befragten bereits eine hohe Kompetenz im Umgang mit Zahlen hatten. Somit ist eine weitere mögliche Erklärung, dass Ärztinnen und

Ärzte aufgrund ihres höheren Bildungsstandes bereits ein besseres Grundverständnis statistischer Zahlen als die allgemeine Bevölkerung haben und nicht von ergänzenden Darstellungen profitieren.

Die Einschätzungen der befragten Hausärztinnen und Hausärzte zu den Früherkennungsmaßnahmen anhand der Zahlen in den Szenarien wichen meist von offiziellen Empfehlungen ab. Im Praxisalltag folgten sie hingegen überwiegend den offiziellen Empfehlungen. Das Vorgehen im Praxisalltag ist offensichtlich nicht unerheblich von den offiziellen Empfehlungen beeinflusst. Für die abweichende Einschätzung anhand der epidemiologischen Zahlen gibt es hingegen verschiedene Erklärungen. Einerseits ist es möglich, dass es den Befragten aufgrund eines geringen statistischen Verständnisses schwer fiel, die Zahlen zu interpretieren. Wie verschiedene Studien zeigten, haben Ärztinnen und Ärzte hier teils erhebliche Defizite [18,19]. Insbesondere die Darstellung in relativen Zahlen im Vergleich zu absoluten Zahlen beeinflusst dabei die Verständlichkeit maßgeblich [18,21,22]. Allerdings waren in den Szenarien stets transparente absolute Risikomaße verwendet worden und es wurden vorrangig relevante Zahlen präsentiert [21]. Eine weitere Erklärung für die Abweichung von offiziellen Empfehlungen ist, dass die Befragten, gerade wegen eines korrekten Verständnisses der Zahlen, Nutzen und Schaden der Früherkennungsmaßnahmen anders bewerten. In einer Studie unter deutschen Hausärztinnen und Hausärzten äußerten sich diese skeptisch bezüglich der Effektivität der bestehenden Früherkennungsmaßnahmen [16]. Es ist nicht verwunderlich, dass die Vorstellungen von einer nützlichen medizinischen Maßnahme verschieden sind. So konnte eine britische Studie zeigen, dass nicht nur PatientInnen und Ärztinnen und Ärzte im Allgemeinen, sondern auch hausärztlich und fachärztlich tätige Ärztinnen und Ärzte unterschiedlich hohe Anforderungen an Nutzen und Schaden von medizinischen Maßnahmen stellen [32]. Die Bewertung von Relevanz anhand von Nutzen und Schaden ist, letztlich auch auf Patientenseite, eine subjektive Entscheidung, die nicht durch die Zahlen, sondern nur auf Basis der Zahlen getroffen werden kann. Dem entsprechend haben Interviews mit Hausärztinnen und Hausärzten gezeigt, dass sie Leitlinien meist skeptisch gegenüberstehen, weil diese zu wenig Bezug zur individuellen Patientin und zum individuellen Patienten hätten [4,5,16].

Limitationen

Starke Einschränkungen der Studie sind der geringe Rücklauf und damit die insgesamt niedrige Anzahl an Befragten. Dies führt

auch zu relativ kleinen Gruppengrößen in den Subgruppenanalysen und eingeschränkter statistischer Power der durchgeführten Tests. Da der Fragebogen online durchgeführt wurde, wurden wahrscheinlich technisch affinere und eher jüngere Teilnehmende erreicht. Das erklärt, neben der Verwendung des Verteilers der Jungen Allgemeinmedizin in Deutschland, warum eine relativ junge Stichprobe selektiert wurde. Die vergleichsweise junge Stichprobe schränkt die Übertragbarkeit auf Hausärztinnen und Hausärzte allgemein ein. Allerdings zeigten sich in den Analysen im Vergleich zwischen jungen und erfahrenen Befragten keine relevanten Unterschiede. Vermutlich haben auch Ärztinnen und Ärzte mit wenig Affinität zu Statistiken seltener an der Untersuchung teilgenommen. Die angeschriebene Gruppe bestand aus jungen Allgemeinmedizinerinnen und -medizinern und Ärztinnen und Ärzten mit engem Kontakt zum allgemeinmedizinischen Institut der Universität. Es ist anzunehmen, dass diese insgesamt aktiver in die Forschung eingebunden sind und möglicherweise die Leitlinienempfehlungen besser kennen, als die Gesamtheit der Hausärztinnen und Hausärzte in Deutschland. Die Eigenschaften der Stichprobe führen also eher zu einer Überschätzung der statistischen Kompetenz und Leitlinienadhärenz der Hausärztinnen und Hausärzte.

Schlussfolgerung

Im Praxisalltag scheint die eigene Bewertung zu Früherkennungsmaßnahmen auf Basis epidemiologischer Zahlen aus der Evidenz nur wenig Bedeutung zu haben. Dies trifft auch bereits auf junge Hausärztinnen und Hausärzte zu. Für eine gemeinsame Entscheidungsfindung mit PatientInnen im Sinne der evidenzbasierten Medizin ist die Nutzung aktueller statistischer Zahlen aber unerlässlich. Hausärztinnen und Hausärzten sollte der Zugang zu verlässlichen, aktuellen Zahlen zu Früherkennungsmaßnahmen erleichtert werden. Vermutlich würden sie sich bei Kenntnis der Zahlen häufiger gegen Früherkennungsmaßnahmen aussprechen.

Erklärungen

Diese Studie wurde ohne spezielle finanzielle Förderung durch öffentliche, kommerzielle oder not-for-profit Einrichtungen durchgeführt. Die vorliegende Arbeit war Teil der Anforderungen zur Erlangung des akademischen Grades „Dr. med.“ an der Medizinischen Fakultät der Friedrich-Alexander Universität Erlangen-Nürnberg.

Interessenkonflikt

Die Autorinnen und Autoren geben an, dass kein Interessenkonflikt besteht. Die ethischen Richtlinien wurden in der Durchführung dieser Studie befolgt. Die Unbedenklichkeitsbescheinigung der Ethik-Kommission der Medizinischen Fakultät der Friedrich-Alexander-Universität Erlangen-Nürnberg wurde eingeholt (327_14c vom 14.11.2014).

Appendix A. Zusätzliche Daten

Zusätzliche Daten verbunden mit diesem Artikel finden sich in der Online-Version unter: [doi:10.1016/j.zefq.2018.06.003](https://doi.org/10.1016/j.zefq.2018.06.003).

Literatur

- [1] (Muster-)Berufsordnung für die in Deutschland tätigen Ärztinnen und Ärzte. http://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/MBO/MBO_02.07.2015.pdf, Bundesärztekammer, 2015, (accessed 23 April 2018).
- [2] Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;312:71–2.
- [3] Grundbegriffe der EbM: Definitionen. <http://www.ebm-netzwerk.de/was-ist-ebm/grundbegriffe/definitionen/>, Deutsches Netzwerk Evidenzbasierte Medizin e.V. (DNEBM), (accessed 05 April 2018).
- [4] Buckler B, Redaelli M, Simic D, Wilm S. „Wir machen doch eigentlich alle EbM!“ – Vorstellungen und Haltungen deutscher Hausärzte zu Evidenzbasierter Medizin und Leitlinien im Praxisalltag: Eine qualitative Studie. *Z. Evid. Fortbild. Qual. Gesundheitswes* 2013;107:410–7.
- [5] Bolter R, Kuhllein T, Ose D, Gotz K, Freund T, Szecsenyi J, Miksch A. Barrieren der Hausärzte gegen Evidenzbasierte Medizin – ein Verständnisproblem? Eine qualitative Studie mit Hausärzten. *Z. Evid. Fortbild. Qual. Gesundheitswes* 2010;104:661–6.
- [6] Holme O, Bretthauer M, Fretheim A, Odgaard-Jensen J, Hoff G. Flexible sigmoidoscopy versus faecal occult blood testing for colorectal cancer screening in asymptomatic individuals. *Cochrane Database Syst. Rev* 2013;CD009259.
- [7] S3-Leitlinie Kolorektales Karzinom, Langversion 1.1. <http://leitlinienprogramm-onkologie.de/Leitlinien.7.0.html>, Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF), 2014, (accessed 11 October 2017).
- [8] Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. *Cochrane Database Syst. Rev* 2013;CD004720.
- [9] Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Berenguer A, Maattanen L, Bangma CH, Aus G, Villers A, Rebillaud X, van der Kwast T, Blijenberg BG, Moss SM, de Koning HJ, Auvinen A, Investigators E. Screening and prostate-cancer mortality in a randomized European study. *N. Engl. J. Med* 2009;360:1320–8.
- [10] T. Kötter, T. Uebel, S1-Handlungsempfehlung: Hausärztliche Beratung zu PSA-Screening. <http://www.degam.de/degam-praxisempfehlungen.html>, Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM), 2013, (accessed 13 October 2014).
- [11] Interdisziplinäre Leitlinie der Qualität S3 zur Früherkennung, Diagnose und Therapie der verschiedenen Stadien des Prostatakarzinoms, Langversion 4.0. <http://leitlinienprogramm-onkologie.de/Prostatakarzinom.58.0.html>, Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF), 2016, (accessed 11 October 2017).
- [12] Pace LE, Keating NL. A systematic assessment of benefits and risks to guide breast cancer screening decisions. *JAMA* 2014;311:1327–35.
- [13] Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. *Cochrane Database Syst. Rev* 2013;CD001877.
- [14] Mammographie-Screening, Früherkennung von Brustkrebs, Was Sie darüber wissen sollten. https://www.krebshilfe.de/fileadmin/Downloads/PDFs/Kampagnen/Broschuere_MammoScreening_2009.pdf, Kooperationsgemeinschaft Mammographie, Deutsches Krebsforschungszentrum – Krebsinformationsdienst, 2009, (accessed 16 Januar 2018).
- [15] Biller-Andorno N, Juni P. Abolishing mammography screening programs? A view from the Swiss Medical Board. *N. Engl. J. Med* 2014;370:1965–7.
- [16] Fisseni G, Golücke A, Abholz H-H. Warum machen deutsche Allgemeinärzte so wenig Früherkennung? – Analyse einer Fokusgruppe. *Z. Allg. Med* 2003;79:591–5.
- [17] S. Falk, M. Reimer, J. Wieschke, S. Heidrich, M. Bogner, Bayerische Mediziner Jahrgang 2015: Rückblick aufs Studium, Weiterbildung und Berufsübergang, Tabellenband. <http://www.bap.ihf.bayern.de/>, Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung (IHF), 2016, (accessed 13 October 2017).
- [18] Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping Doctors and Patients Make Sense of Health Statistics. *Psychol. Sci. Public Interest* 2007;8:53–96.
- [19] Wegwarth O, Schwartz LM, Woloshin S, Gaissmaier W, Gigerenzer G. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Ann. Intern. Med* 2012;156:340–9.
- [20] Gigerenzer G, Edwards A. Simple tools for understanding risks: from innumeracy to insight. *BMJ* 2003;327:741–4.
- [21] J. Lühnen, M. Albrecht, I. Mühlhauser, A. Steckelberg, Leitlinie evidenzbasierte Gesundheitsinformation. <http://www.leitlinie-gesundheitsinformation.de/wp-content/uploads/2017/07/Leitlinie-evidenzbasierte-Gesundheitsinformation.pdf>, 2017, (accessed 14 April 2018).
- [22] Covey J. A meta-analysis of the effects of presenting treatment benefits in different formats. *Med. Decis. Making* 2007;27:638–54.
- [23] Woloshin S, Schwartz LM. Communicating data about the benefits and harms of treatment: a randomized trial. *Ann. Intern. Med* 2011;155:87–96.
- [24] Berry DC, Knapp P, Raynor T. Expressing medicine side effects: assessing the effectiveness of absolute risk, relative risk, and number needed to harm, and the provision of baseline risk information. *Patient Educ. Couns* 2006;63:89–96.
- [25] Rajkumar SV, Sampathkumar P, Gustafson AB. Number needed to treat is a simple measure of treatment efficacy for clinicians. *J. Gen. Intern. Med* 1996;11:357–9.
- [26] Galesic M, Garcia-Retamero R, Gigerenzer G. Using icon arrays to communicate medical risks: overcoming low numeracy. *Health Psychol* 2009;28:210–6.
- [27] Bell DS, Sobolevsky S, Day FC, Hoffman JR, Higa JK, Wilkes MS. The stadium diagram, a web-based tool for visualizing the expected outcomes of alternative clinical management strategies. *AMIA Annu Symp Proc* 2005;36–40.
- [28] Krebs in Deutschland für 2013/2014, 11. Ausgabe. https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/krebs_in_deutschland_inhalt.html, Robert Koch-Institut, Gesellschaft der

- epidemiologischen Krebsregister in Deutschland e.V., 2017, (accessed 16 January 2018).
- [29] Ilic D, Murphy K, Green S. What do general practitioners think and do about prostate cancer screening in Australia? *Aust. Fam. Physician* 2013;42:904–8.
- [30] Ulvenes LV, Aasland O, Nylenna M, Kristiansen IS. Norwegian physicians' knowledge of and opinions about evidence-based medicine: cross-sectional study. *PLoS One* 2009;4:e7828.
- [31] Diener A, Celemin-Heinrich S, Wegscheider K, Kolpatzik K, Tomaschko K, Altner A, Donner-Banzhoff N, Haasenritter J. In-vivo-validation of a cardiovascular risk prediction tool: the arriba-pro study. *BMC Fam. Pract* 2013; 14:13.
- [32] Steel N. Thresholds for taking antihypertensive drugs in different professional and lay groups: questionnaire survey. *BMJ* 2000;320:1446–7.
- [33] Versorgungsatlas Hausärzte, Darstellung der regionalen Versorgungssituation sowie der Altersstruktur in Bayern, Januar 2018. https://www.kvb.de/fileadmin/kvb/dokumente/UeberUns/Versorgung/KVB-Versorgungsatlas_Hausaerzte.pdf, Kassenärztliche Vereinigung Bayerns, 2018, (accessed 19 April 2018).