

1 Prinzipien und Methoden von Früherkennungsuntersuchungen

Bernt-Peter Robra und Norbert Schmacke

C. Günster | J. Klauber | B.-P. Robra | N. Schmacke | C. Schmuker (Hrsg.) Versorgungs-Report Früherkennung
DOI 10.32745/9783954664023-1, © MWV Medizinisch Wissenschaftliche Verlagsgesellschaft Berlin 2019

Abstract

Früherkennung von Erkrankungen ist nur dann legitim, wenn durch hinreichend große randomisierte kontrollierte Studien belegt ist, dass sie tatsächlich einen klinisch relevant günstigeren Verlauf der Erkrankung bewirkt. Eine naive Gleichsetzung von „früh(er)“ mit „wirksam(er)“ und „kostengünstig(er)“ ist nicht haltbar. In dem Umfang, in dem Primärprävention möglich ist, z.B. durch Verzicht auf Noxen, oder kurative Behandlung erfolgreich ist, verliert Früherkennung an Bedeutung.

Der vorliegende Beitrag beschreibt Begriffe und Prinzipien, die bei Vorbereitung, Einführung und Evaluation von Früherkennungsprogrammen berücksichtigt werden müssen. Selbst bei Vorliegen hochwertiger Evidenz aus (internationalen) Studien verlangt die Entscheidung, ein Früherkennungsprogramm flächendeckend einzurichten, mutige Verallgemeinerungen auf die Funktionsfähigkeit des eigenen Gesundheitswesens. Das unterstreicht den Stellenwert einer begleitenden formativen Evaluation. Die Möglichkeiten dazu haben sich in Deutschland schrittweise verbessert. Zu wesentlichen Fragen, nicht zuletzt zum Stellenwert des Screenings auf Haut- und Prostatakrebs und der Früherkennung in der Primärversorgung, gibt es aber inter-

nationaler Klärungsbedarf, zu dem eine Beobachtung laufender Aktivitäten der GKV nur indirekt beitragen kann.

Welche Information darf ein individueller Versicherter daher in Sachen Screening erwarten, welche kann er bekommen und welche erhält er? Standards einer Gesundheitskompetenz der Bürger fördernden Informationspolitik werden zunehmend erkennbar. Doch ohne Polemik ist festzustellen: Der Nutzen selbst gut begründeter und durchgeführter Screening-Programme wird i.d.R. in öffentlich verbreiteten Aufklärungsmaterialien und in der individuellen Arzt-Patienten-Kommunikation deutlich überschätzt. Das große Thema „paternalistische vs. versichertenorientierte Aufklärung“ ist kein Thema allein der Patientenrechte. Im Bereich der Früherkennung ist es auch eine Frage der Rechte der Bürger.

Early detection of diseases requires evidence of clinically relevant improvements in the natural course of the disease, generally provided by sufficiently powered randomized controlled trials. A simplistic equation of „earlier“ with „more effective“ or „more efficient“ is not tenable. Early detection loses relevance to the extent that primary prevention, e.g. by avoiding hazards, or curative interventions become possible.

This introduction presents terminology and principles which must be considered in preparation, implementation, and evaluation of screening programmes. Even if high grade (international) evidence is available the decision to implement a population-based screening programme requires some courageous extrapolations towards the functional capability of the targeted health system. This points to the significance of concomitant formative evaluation. The evaluative potential in the German health system has gradually improved. Major clarifications, however, are necessary in the international realm, e.g. concerning the effectiveness of screening for skin or prostate cancer and the working of multiphasic screening in primary care. Here monitoring of ongoing health services in Germany can contribute, albeit marginally.

Which information regarding screening measures may an individual expect, which could he possibly get and which will eventually be available to him? Standards of communication fostering health literacy of the public are becoming available. But it can be stated without being polemic that the benefits of screening programmes, even those based on evidence and conducted well, is grossly overvalued in publicly disseminated information materials and in doctor-patient-encounters. The major topic „paternalistic vs. shared decision making“ concerns not alone patients' rights. In screening, this is also a question of the citizens' rights.

1.1 Früherkennung in der gesetzlichen Krankenversicherung

Der Gesetzgeber definiert Umfang und Qualität der Krankenversorgung, die den Versicherten zusteht und die Krankenkassen und Leistungserbringer sicherzustellen haben, durch offene Rechtsbegriffe (ausreichend, wirtschaftlich, zweckmäßig, notwendig: § 12 SGB V). Zu Früherkennungsprogrammen dagegen hat er im vierten Abschnitt des Sozialgesetzbuchs V Vorgaben gemacht (§§ 25, 25a und 26 SGB V). Der Gemeinsame Bundesausschuss (G-BA) legt mit seiner Richtlinienkompetenz nach § 92 SGB V das Nähere über Früherkennungsmaßnahmen, organisierte Früherkennungsprogramme und die Qualitätssicherung der Früherkennungsuntersuchungen fest:

- Richtlinie zur Empfängnisregelung und zum Schwangerschaftsabbruch (vormals Sonstige-Hilfen-Richtlinie) mit der Vorgabe eines jährlichen genitalen Chlamydien-Screenings bei Frauen bis zum abgeschlossenen 25. Lebensjahr
- Richtlinien über die ärztliche Betreuung während der Schwangerschaft und nach der Entbindung (Mutterschafts-Richtlinien)
- Richtlinie über die Früherkennung von Krankheiten bei Kindern (Kinder-Richtlinie)
- Richtlinien über die Früherkennungsuntersuchungen auf Zahn-, Mund- und Kieferkrankheiten (Zahnärztliche Früherkennung)
- Richtlinie zur Jugendgesundheitsuntersuchung
- Richtlinie über die Gesundheitsuntersuchungen zur Früherkennung von Krankheiten (Gesundheitsuntersuchung-Richtlinie)
- Richtlinie über die Früherkennung von Krebserkrankungen (Krebsfrüherkennungsrichtlinie)
- Richtlinie für organisierte Krebsfrüherkennungsprogramme (oKFE-RL)
- Richtlinie Ultraschall-Screening auf Bauch-aortenaneurysmen (US-BAA-RL)

Das im Einheitlichen Bewertungsmaßstab geregelte sog. hausärztlich-geriatrische Basisassessment ist eine weitere komplexe Früherkennungsmaßnahme, allerdings ohne Richtlinie des G-BA. Insgesamt hält die gesetzliche Krankenversicherung also ein lebensbegleitendes Früherkennungsangebot vor (siehe dazu auch die Übersicht des Gemeinsamen Bundesausschusses: <https://www.g-ba.de/institution/themenschwerpunkte/frueherkennung/ueberblick/>).

Wie sind diese Regulierungen der sog. Selbstverwaltung im Gesundheitswesen in den Stand des Wissens um Früherkennungsprinzipien und -methoden einzuordnen? Ein Blick auf die Früherkennungsmaßnahmen anderer entwickelter Länder zeigt: Zielkrankheiten, Einladungsverfahren, eingesetzte Suchtests und die weitere Abklärungsdiagnostik, der Doku-

mentationsaufwand und – nicht leicht zu beurteilen – wahrscheinlich auch die erreichten Ergebnisse und damit die Wirtschaftlichkeit sind international unterschiedlich. So verzichtet z.B. selbst das sonnenverwöhnte Australien auf die Früherkennung von Hautkrebs (Cancer Council Australia's National Skin Cancer Committee 2014) und Schweizer Kantone schaffen die Einladung zum Mammografie-Screening wieder ab (Föry 2017), nachdem das unabhängige Swiss Medical Board sich wegen eines konstatierten ungünstigen Nutzen-Risiko-Verhältnisses und eines ungünstigen Kosten-Wirksamkeits-Verhältnisses gegen das systematische Mammografie-Screening ausgesprochen hatte (Fachgremium Swiss Medical Board 2013). Internationale Vergleiche verlangen also eine kritische Überprüfung der deutschen Programme mit dem Ziel einer Verbesserung von Angebot, Organisation und Durchführung.

Selbst innerhalb Deutschlands gibt es erkennbar unterschiedliche Bewertungen des Notwendigen, Zweckmäßigen und Ausreichenden. So haben Vertreter der Deutschen Gesellschaft für Allgemeinmedizin (DEGAM) abweichende Voten in die Leitlinien der Fachgesellschaften zur Früherkennung von Hautkrebs (Chenot u. Egidi 2015) und von Prostatakarzinomen (AWMF Leitlinienprogramm Onkologie 2018; Kötter 2016) eingebracht, unter anderem mit dem Ziel, potenzielle Früherkennungsteilnehmer vor ungerechtfertigten Erwartungen und einer ungünstigen Nutzen-Risiko-Bilanz zu schützen.

Aus der Perspektive der bevölkerungsbezogenen Versorgungsforschung gibt es – wie die einzelnen Beiträge dieses Versorgungs-Reports zeigen – zahlreiche offene Fragen zur evidenzbasierten Begründung laufender Früherkennungsprogramme, zu ihrer patientenorientierten Durchführung und zur methodischen Bewertung, damit auch zur weiteren Entwicklung. Zum Beispiel fehlt dem im hausärztlich-geriatrischen Basisassessment enthaltenen Demenz-Screening die Evidenzbasis (s. Kap. 11). Darüber hinaus können praktisch täglich neue

Erkenntnisse der klinischen und der biomedizinischen Forschung potenziell auch zur Früherkennung von Krankheiten beitragen, weil sie ein besseres Verständnis biologischer Mechanismen, verbesserte diagnostische Verfahren oder wirksamere therapeutische Strategien zum Gegenstand haben. Diese Potenziale bedürfen allerdings erst einer aufwändigen kontrollierten Erprobung in symptomfreien Versichertengruppen, bevor sie flächendeckend und qualitätsgesichert umgesetzt werden können. Derartige grundlegende Erprobungen, insbesondere randomisierte kontrollierte Studien (RCTs, Goldstandard-Studien), fehlen für den Beginn des deutschen Krebsfrüherkennungsprogramms, was bis heute zu Unsicherheiten seiner Bewertung beiträgt und bei in Aussicht genommenen Modifikationen und Ergänzungen zu versorgungswissenschaftlichen, ethischen und rechtlichen Diskursen beiträgt.

Die folgenden Abschnitte dieses Beitrags sind wie folgt gegliedert: Zunächst werden grundsätzliche Prinzipien und Begriffe der Früherkennung geklärt und krankheitsübergreifend methodische Grundlagen der Bewertung von Früherkennungsprogrammen erläutert. Dabei wird unterschieden zwischen der Perspektive eines Individuums, das vor der Frage steht, an einem Früherkennungsprogramm für symptomlose Personen teilzunehmen (Individualperspektive), und der Perspektive der Entscheidungsverantwortlichen für Strukturen, Prozesse, Qualitätssicherung und Finanzierung im Gesundheitswesen (gesellschaftliche Perspektive). Weiter wird unterschieden zwischen der Perspektive vor Einführung eines systematischen Screenings (ex ante) und der Perspektive der Bewertung eines laufenden Screening-Programms (ex post). In der Situation ex ante sind die individuelle und die gesellschaftliche Perspektive strukturell nicht zu trennen. Sie können sich aber in der Bewertung von Risiken und Nutzen einzelner Gesundheitszustände unterscheiden. Die Abklärung von Verdachtsfällen und selbst invasive Eingriffe werden vorhersehbar Untergruppen der Teilnehmer betreffen, die

selber davon keinen Nutzen, sondern nur Nachteile haben werden. Daher muss ein symptomloser potenzieller Teilnehmer schon vor einer Entscheidung über die Teilnahme an einem wohldefinierten Früherkennungsprogramm umfassend über alle Folgen seiner Teilnahme aufgeklärt werden. Abschnitt 1.3 untersucht, wie Bürger in die Lage versetzt werden können, solche Entscheidungen selbstbestimmt, d.h. ggf. auch gegen den Hauptstrom der gesellschaftlichen Meinung und den Paternalismus in der Medizin, zu treffen.

1.2 Prinzipien und Begriffe

Früherkennungsuntersuchungen, auch Filteruntersuchungen oder im internationalen Sprachgebrauch Screening genannt, richten sich definitionsgemäß an symptomlose Zielpersonen. Sobald ein Versicherter sich zur Abklärung von (selbst bemerkten) Symptomen an einen Arzt wendet, handelt es sich um ein Problem der Krankenversorgung. Dieser Unterschied ist von praktischer und ethischer Bedeutung. Denn wenn ein Versicherter ärztlichen Rat und Hilfe sucht, hat er die Initiative und muss nach Aufklärung (§ 630e BGB) selbst entscheiden, ob er die ärztlicherseits vorgeschlagenen Optionen akzeptiert oder nicht. Wenn aber ein ganzes Gesundheitssystem, eine Krankenversicherung oder ein Versorger symptomlosen Personen Untersuchungsverfahren aktiv anbietet oder die Inanspruchnahme nahelegt, müssen die Anbieter vorher allgemeinverständlich darlegen, dass für jeden potenziellen Teilnehmer und für die Zielgruppe insgesamt eine hinreichende Wahrscheinlichkeit für einen nicht trivialen medizinischen (Netto-) Nutzen besteht und welche unerwünschten Wirkungen gegen eine Inanspruchnahme sprechen.

An dieser Beweislast ändert sich nichts dadurch, dass das Ergebnis eines Screening-Tests (noch) nicht diagnostisch belastbar ist. Sein Zweck ist, die Wahrscheinlichkeit für das Vor-

liegen einer ernsthaften und wirksam behandelbaren Zielkrankheit so weit zu erhöhen, dass Abklärungsuntersuchungen gerechtfertigt sind und vorverlegte Therapiemaßnahmen für den Teilnehmer eine bessere Prognose erwarten lassen als die übliche Versorgung ohne Screening.¹

Früherkennung wird herkömmlich als „sekundäre Prävention“ der Primärprävention an die Seite gestellt. Primärprävention soll das Auftreten von Krankheiten verhindern, weniger wahrscheinlich machen oder verzögern (Walter et al. 2012, S. 196). Früherkennung verhütet keine Krankheit, sondern soll Langzeitergebnisse der Versorgung verbessern. Insofern ist der Begriff „sekundäre Prävention“ potenziell missverständlich. Früherkennung steht der Rehabilitation näher als der Primärprävention.

Wenn wir abgekürzt von „Screening-Programmen“ sprechen, ist ein mehrstufiges Vorgehen gemeint:

- a) die Auswahl einer geeigneten Zielkrankheit, eines zu ihrer Erkennung geeigneten Früherkennungstests und einer umschriebenen symptomlosen Zielgruppe
- b) die Einladung zur Teilnahme an diesem Früherkennungsprogramm
- c) die aufgeklärte Entscheidung einer symptomlosen Person zur Teilnahme
- d) die Durchführung notwendiger Folgeuntersuchungen zur Abklärung eines auffälligen Früherkennungstests
- e) die indizierte Frühintervention
- f) die begleitende Dokumentation, Erfolgssicherung und Bewertung (Evaluation) des Vorgehens

1 Das Oxford Textbook of Medicine definiert: „Medical screening is the systematic application of a test or inquiry to identify individuals at sufficient risk of a specific disorder to benefit from further investigation or direct preventive action (these individuals not having sought medical attention on account of symptoms of that disorder). Key to this definition is that the early detection of disease is not an end in itself; bringing forward a diagnosis without altering the prognosis is useless and may be harmful.“ (<http://oxfordmedicine.com/view/10.1093/med/9780199204854.001.1/med-9780199204854-chapter-030302>)

Früherkennung ist also kein Selbstzweck („je früher desto besser“), sondern erst gerechtfertigt, wenn sich am Ende eines komplexen Prozesses die gesundheitliche Perspektive von Menschen spürbar verbessern lässt. Der Nachweis einer solchen Verbesserung setzt kontrollierte Langzeitstudien voraus, die schwer zu organisieren und durchzuhalten sind und in Deutschland fehlen (siehe dazu die einzelnen Beiträge in diesem Band). Die genannten Stufen des Vorgehens strukturieren die vorbereitende und die nachgehende Bewertung eines Früherkennungsprogramms. Sie werden im Folgenden vertieft.

Zu a) die Auswahl einer geeigneten Zielkrankheit, eines zu ihrer Erkennung geeigneten Früherkennungstests und einer umschriebenen symptomlosen Zielgruppe

Unter den Zielkrankheiten laufender Screening-Programme sind die Karzinome von Gebärmutterhals, Brust und Dickdarm besonders prominent. Eine Teilnahme daran wird u.a. vom „European Code against Cancer“ unterstützt (Armaroli et al. 2015). Haut- und Prostatakrebs finden weniger Unterstützung. Die systematische Suche nach unentdeckten Neuroblastomen im Kindesalter hat sich nach Evaluation als nicht wirksam erwiesen (Schilling et al. 2002; Woods et al. 2002). Weitere Zielkrankheiten, darunter fachlich umstrittene oder nicht empfehlenswerte, sind Gegenstand weiterer Kapitel dieses Bandes.

Als Geburtsstunde heutiger Gesunden-Untersuchung kann die Suchen nach psychiatrischen Erkrankungen bei Musterungen der USA-Armee im Jahr 1917 gelten (Morabia u. Zhang 2004). Musterungsuntersuchungen hatte es natürlich vorher schon gegeben, doch tauchte in diesem Rahmen erstmals die Frage nach der Validität der Suchverfahren, den falsch positiven und falsch negativen Probanden auf. Ziel war allerdings nicht Frühtherapie, sondern eine Eignungsuntersuchung. Mit der zunehmenden

Verfügbarkeit von Laborautomaten richteten sich Hoffnungen auf ein multimodales Screening mit Kombination von Anamnese, körperlicher Untersuchung, Laborwerten und Gesundheitsberatung. Sie haben sich nicht erfüllt (z.B. South East London Screening Study 1977; Caley et al. 2014; Jørgensen et al. 2014), der deutsche Check-up 35 ist umstritten und revisionsbedürftig (siehe auch Mühlhauser 2014 und Kap. 9). Der Einsatz von EKG und Belastungs-EKG zur Früherkennung von Herzerkrankungen will selbst bei Menschen mit deutlich erhöhtem Risiko für eine koronare Herzerkrankung gut überlegt sein (Bhatia u. Dorian 2018). Auch übertragbare Krankheiten sind Gegenstand systematischer Suchtests, z.B. Syphilis, HIV/AIDS, Röteln oder MRSA. Screening darauf unterscheidet sich von dem auf andere Zielkrankheiten ethisch und ökonomisch dadurch, dass es unter Umständen Nutzen für Dritte stiften kann, die bisher nicht betroffen sind. Trotzdem wird in Deutschland auf eine Röntgen-Reihenuntersuchung der Wohnbevölkerung auf Tuberkulose, viele Jahre Teil öffentlicher Gesundheitspflege, verzichtet – Folge einer Risiko-Nutzen-Abwägung, die bei Migranten aus Ländern mit hohem Vorkommen an Tbc anders ausfällt.

Wilson und Junger (1968, S. 26–27) haben die weitreichenden Vorbedingungen dieses definitorischen Schritts wie folgt zusammengefasst (vgl. auch § 25 Abs. 3 SGB V):

1. The condition sought should be an important health problem.
2. There should be an accepted treatment for patients with recognised disease.
3. Facilities for diagnosis and treatment should be available.
4. There should be a recognisable latent or early symptomatic stage.
5. There should be a suitable test or examination.
6. The test should be acceptable to the population.
7. The natural history of the condition, including development from latent to declared disease, should be adequately understood.

8. There should be an agreed policy on whom to treat as patients.
9. The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole.
10. Case-finding should be a continuing process and not a 'once and for all' project.

Für die Auswahl eines zur Früherkennung geeigneten Tests sind neben Praktikabilität und Kosten vor allem Parameter seiner Eignung (Gültigkeit, Validität) relevant: Sensitivität und Spezifität. Die Sensitivität Se gibt die Wahrscheinlichkeit an, dass eine vorhandene Zielkrankheit tatsächlich erkannt wird. Die Spezifität Sp gibt die Wahrscheinlichkeit an, dass eine Person ohne Zielkrankheit als unauffällig eingestuft wird.

Das Komplement der Sensitivität ($1-Se$), die Wahrscheinlichkeit, eine vorhandene Ziel-läsion zu übersehen, ist ein bedeutsames Problem für Früherkennungsteilnehmer und Leistungserbringer, das möglichst klein gehalten werden sollte. Aber auch die Unspezifität ($1-Sp$), die Wahrscheinlichkeit eines falsch positiven Testergebnisses bei Teilnehmern ohne Zielläsion, führt zu unerwünschten medizinischen Folgen. Sie sind nicht zuletzt deswegen bedeutsam, weil die Teilnehmer ohne Zielläsion die große Mehrheit aller Teilnehmer stellen. Nach 10 Jahren jährlicher Screening-Mammografie wird z.B. mehr als die Hälfte aller teilnehmenden Frauen wenigstens eine falsch positive Nachuntersuchung erhalten, 7% bis 9% eine falsch positive Biopsieempfehlung. Ein Screening nur alle zwei Jahre reduziert diese kumulative Wahrscheinlichkeit, erhöht aber die Wahrscheinlichkeit einer Diagnose erst in einem späten Stadium (Hubbard et al. 2011).

Sensitivität und Spezifität eines Tests hängen typischerweise gegenläufig miteinander zusammen. Die Wahl des Grenzwertes, von dem ab ein Testergebnis als auffällig beurteilt

wird und deswegen zu Konsequenzen führt, beeinflusst beide Testgüte-Parameter, damit Aufwand, Kosten und Erfolg des ganzen Programms. Das „richtige“ Verhältnis von Sensitivität und Spezifität hängt nicht zuletzt davon ab, wie die Regulatoren oder die Zielgruppe das Verhältnis des Nutzens richtig erkannter Zielkrankheiten zum Schaden durch Testaufwand, Abklärungsaufwand und falsch positive, d.h. schädliche Folgeinterventionen bewerten. Ein schmerzhaftes Beispiel ist die notwendige Inkaufnahme des Todes gesunder Föten bei der systematischen Früherkennung von Fehlbildungen in der Schwangerschaft (Felder et al. 2003; Felder u. Robra 2006). Erstmals zur Verwendung immunologischer Tests auf Blut im Stuhl (iFOBT) hat der Gemeinsame Bundesausschuss explizit Mindestgrenzen für Sensitivität und Spezifität vorgegeben (Bundesministerium für Gesundheit 2016). Für die Entdeckung von Darmkrebsen und fortgeschrittenen Polypen muss in nur einer Stuhlprobe eine Sensitivität von mindestens 25% bei einer Spezifität von mindestens 90% erreicht werden. Referenztest ist das Ergebnis einer Koloskopie. Mit diesen zurückhaltenden Mindestvorgaben offenbart der G-BA implizit eine Güterabwägung von Nutzen und Schäden im Rahmen des Darmkrebsfrüherkennungsprogramms.

Ein weiteres, unvermeidbares Problem ist als Überdiagnostik bekannt. Es bedeutet, dass in Bevölkerungen mit Früherkennungsprogramm für einen bestimmten Anteil der entdeckten und dann auch behandelten Fälle die Zielkrankheit zu Lebzeiten nicht mehr klinisch manifest oder gar lebenszeitbegrenzend geworden wäre. Es handelt sich um ein bevölkerungsbezogenes Konzept, d.h. individualmedizinisch tätige Kliniker und Pathologen können es nicht erkennen – sie sehen und behandeln einen nachgewiesenen Fall. Überdiagnostik trifft vor allem bei Zielkrankheiten zu, die bevorzugt im höheren Alter auftreten. Sie kann in Studien quantifiziert werden, die Neuerkrankungsziffern und Sterblichkeit in Bevölkerungen mit und ohne Möglichkeit zur Teil-

Tab. 1 Erwarteter Nutzen eines Screening-Programms – Ergebnis des Screening-Tests, Vorliegen der Krankheit und Bewertungen des Nettonutzens der vier Zustände

		Krankheit	
		ja	nein
Test	positiv	$Pr \cdot Se \cdot NU_{(T+K+)}$	$(1 - Pr) \cdot (1 - Sp) \cdot NU_{(T+K-)}$
	negativ	$Pr \cdot (1 - Se) \cdot NU_{(T-K+)}$	$(1 - Pr) \cdot Sp \cdot NU_{(T-K-)}$

Pr ist die Prävalenz der Zielläsionen, bei einem einzelnen Teilnehmer sein Krankheitsrisiko a priori; Se Sensitivität und Sp Spezifität sind Kennwerte der Testgüte; NU steht für Nettonutzen (net utility), der Nettonutzen der vier Zustände wird durch Indizes unterschieden: K+ bzw. K- signalisieren den wahren Gesundheitszustand, T+ bzw. T- das Testergebnis

nahme am Früherkennungsprogramm miteinander vergleichen. Der Anteil überdiagnostizierter Fälle an allen Fällen ist z.B. hoch beim Screening auf Prostatakarzinom (Robra et al. 2013), aber auch beim Brustkrebs-Screening nicht zu vernachlässigen (Autier et al. 2017) und reduziert natürlich den Nutzen der Früherkennung. Jedes Absenken diagnostischer Schwellen verstärkt Überdiagnostik (Welch et al. 2011). Forschung zu aussagefähigen Progressionsmarkern könnte helfen, unnötige Interventionen zu reduzieren.

Falsch positive und falsch negative Testergebnisse und Überdiagnostik sind nicht nur Herausforderungen der individuellen Teilnahmeaufklärung, sondern Probleme für die Konzeption eines Früherkennungsprogramms ex ante. Denn das Angebot eines Früherkennungsprogramms bedeutet de facto die Einladung zur Teilnahme an einer Lotterie: Einige Teilnehmer werden einen Nutzen realisieren können, andere werden Schäden davontragen. Es handelt sich also um einen Verstoß gegen das Gebot, nicht zu schaden (Non-Malefizienz). Die hinter dem Programmangebot stehende Güterabwägung trifft der Regulator, noch bevor ein potenzieller Teilnehmer individuelle Präferenzen geltend machen kann.

Tabelle 1 illustriert vier elementare Zustände im Screening: Der Screening-Test kann positiv oder negativ ausfallen, die Krankheit liegt vor oder nicht. Zunächst vernachlässigen wir die dritten Multiplikatoren, die Nutzenparameter.

Die Wahrscheinlichkeit, einen der vier Zustände zu erreichen, ist eine Funktion von Prävalenz und Testgüte. Die Wahrscheinlichkeit eines Teilnehmers, sich im Feld links oben zu finden, ist gegeben durch das Produkt aus Prävalenz und Sensitivität. Das Feld darunter quantifiziert die Wahrscheinlichkeit eines falsch negativen Ausgangs, das Feld rechts oben die Wahrscheinlichkeit eines falsch positiven. Das Feld rechts unten beschreibt die Wahrscheinlichkeit, zur Gruppe der negativ getesteten Gesunden zu gehören. Das ist typischerweise die größte Gruppe.

Der dritte Parameter NU in jeder Zelle steht für den Nettonutzen, der jenem Zustand zugeschrieben wird. $NU_{(T+K+)}$ steht für den Nettonutzen richtig erkannter Kranker. Nettonutzen heißt, dass der medizinisch bewertete Nutzen dieses Zustands, z.B. ein Gewinn an Lebensdauer, schon um unerwünschte Effekte vermindert wurde, z.B. um einen Verlust an Lebensqualität relativ zum Status vor Testung. $NU_{(T+K+)}$ muss positiv sein, sonst macht das Screening keinen Sinn. $NU_{(T+K-)}$ steht für den Nettonutzen, mit dem ein falsch positiver Testausgang bewertet wird. Er ist typischerweise negativ. Je nach Perspektive kann die Nutzenbewertung zusätzlich direkte und indirekte Kosten einbeziehen und Nutzenströme über die Zeit diskontieren.

Die Summe der vier bewerteten Zustände ist der erwartete Gesamtnutzen von Früherkennungsprogrammen auf der Ebene der symp-

tomlosen Zielbevölkerung, aber auch für ein symptomloses Individuum in Unkenntnis der es erwartenden Zustände. Die individuelle und die gesellschaftliche Bilanz gehen also von dem gleichen (evidenzgestützten) Mengengerüst aus (Prävalenz bzw. Risiko, Testgüte). Auf beiden Ebenen muss die Nutzenbilanz positiv sein, um ein Screening-Programm überhaupt anbieten zu können bzw. eine Teilnahme individuell in Erwägung zu ziehen. Allerdings kann sich die durch den Regulator unterstellte „gesellschaftliche“ Bewertung der Zustände von einer individuellen Bewertung unterscheiden. Das unterstreicht die Notwendigkeit der individuell aufgeklärten Zustimmung auch bei „programmierten“ Screening-Angeboten.

Ein Individuum kann für sich selber positive und negative Effekte miteinander verrechnen. Solche individuellen Abwägungen verschieden dimensionierter Effekte finden überall im Gesundheitswesen statt, z. B. wenn die Einnahme schmerzlindernder und entzündungshemmender nichtsteroidaler Antirheumatika (NSAR) gegen eine damit verbundene erhöhte Blutungsneigung abgewogen werden muss. Damit die Abwägung auch beim Screening möglich ist, braucht das Individuum allerdings nicht die abstrakte Erwartungsnutzentheorie der Tabelle 1, sondern für ihn verständliche Quantifizierungen der zu erwartenden Versorgungspfade (z. B. Raftery u. Chorozeoglou 2011; Welch u. Passow 2014): Welchen Aufwand bedeutet z. B. eine Darmspiegelung für ihn, wie viele von 1.000 Teilnehmern werden zur Abklärung auffälliger Befunde einbestellt, was kann dabei passieren, welche Behandlungsoptionen stehen zur Verfügung? Und wie viele von 1.000 Teilnehmern werden mittel- und langfristig welchen Nutzen erreichen („number needed to screen“), wie viele werden aber auf welche Weise falsch positiv versorgt oder überdiagnostiziert?

Individuelle Zustimmung zur Teilnahme löst jedoch auf der gesellschaftlichen Ebene das ethische Dilemma nicht, dass ein Programmangebot im Interesse eines positiven Nettonutzens eines Teils der Teilnehmer einem anderen

Teil der Teilnehmer einen negativen Nettonutzen, d. h. Schaden, zumutet. Regulatoren schwächen dieses Dilemma für sich zunehmend dadurch ab, dass sie in einem mittleren Risikobereich wohlstrukturierte Programmvorgaben vermeiden, sondern die Entscheidung über die Früherkennungsteilnahme explizit den Versicherten und ihren Ärzten anheimstellen, die das Vorgehen risiko- und präferenzabhängig kontextuieren, wie sie es in der klinisch-kurativen Individualmedizin kennen („discretionary care“). Bei Früherkennungsmaßnahmen im Niedrigprävalenzbereich und mit großer Unsicherheit der Nutzenbewertung müssen sie jedoch kritischer entscheiden als bei akut-dringlichen „kurativen“ Maßnahmen, d. h. Ärzte wie Patienten müssen die komplexe Evidenzlage angemessen berücksichtigen können. Ärztliche Einwilligungsaufklärung muss im Übrigen in einer Situation mit wenig Handlungsdruck noch gründlicher sein als in akuten klinischen Situationen (Windus et al. 1992).

Eine andere Folge der Entscheidungskomplexität sind „politische Ersatzvornahmen“. Beim Mammografie-Screening kommt dies besonders deutlich zum Ausdruck: Hier entschied das Parlament nach öffentlicher Anhörung von Sachverständigen einmütig, dass es einen großen Screening-Nutzen gebe und Deutschland gegenüber den europäischen Leitlinien im Vorzug sei (Deutscher Bundestag 2002).

So wie Mengengerüst (geringe Prävalenz der Zielkrankheiten!), Testgüte und Nutzen-Relationen von Früherkennungsprogrammen typischerweise ausfallen, ist ihr gesellschaftlicher Gesamtnutzen sensitiv gegenüber der Höhe der direkten Durchführungskosten und den unerwünschten (falsch positiven) Wirkungen für die Gesunden. Der in der Werbung für das Programm herausgestellte Nutzen für früh erkannte Fälle beeinflusst das Gesamtergebnis eher wenig, da diese Fälle zwar großen Nutzen gewinnen können, aber selten sind. Dies spricht dafür, ein Screening-Programm schonend und kostensensibel zu organisieren. Für

die Versicherten der gesetzlichen Krankenversicherung sind die direkten Kosten zudem durch das Sachleistungsprinzip auf einen Wert von Null heruntersubventioniert. Für Ärzte bedeuten Früherkennungsprogramme wenig risikante Einnahmen. Diese Aspekte tragen zu Bewertungsunterschieden zwischen Systemebene und Individualebene bei, obwohl eine Früherkennungsteilnahme auch für individuelle Teilnehmer mit Opportunitätskosten verbunden ist.

Als weiteres Problem der Nutzenbewertung können bei zunehmend perfektionierten Untersuchungsmethoden deutlich häufiger als früher Frühstadien ans Licht gebracht werden, von denen unbekannt ist, wann oder ob sie in klinisch relevante Stadien umschlagen werden und ob sie mit den für fortgeschrittene Stadien entwickelten Therapieschemata wirksam oder schonend genug behandelt werden können. Die Verschiebung im Spektrum entdeckter Stadien bringt also neue Krankheitsentitäten hervor, die aus eigenem Recht prognostisch und mit kontrollierten Therapiestudien eingeordnet werden müssen. Das traf z.B. auf das sog. ductale Carcinoma in situ (DCIS) der Brust zu, dessen Häufigkeit bei systematischer Suche per Mammografie nennenswert zunahm (Jatoi u. Baum 1995; Correa et al. 2010). Das trifft auch für das lokal begrenzte Prostatakarzinom zu, für dessen Management kontrollierte Überwachung hinreichen kann (Bill-Axelsson et al. 2014; Hamdy et al. 2016; Wilt et al. 2017; Wellek et al. 2017). Damit wird jedoch die Grundidee der Früherkennung, die Vorverlegung des Diagnosezeitpunktes, hinfällig.

Zu b) die Einladung zur Teilnahme an diesem Früherkennungsprogramm

Als Zielgruppe für die Einladung zum Test kommt nur eine Versichertengruppe in Betracht, für die der Test akzeptabel ist (Tests, die auf Schwangerschaftsunterbrechung hinauslaufen, sind z.B. nicht in allen Bevölkerungs-

gruppen akzeptabel), für die die Zielkrankheit hinreichend häufig oder belastend ist² (das kann auf eine untere Altersgrenze zum Screening oder eine andere Abgrenzung einer Gruppe hohen Risikos hinauslaufen) und für die ein möglicher Früherkennungsnutzen sich noch zu Lebzeiten manifestieren kann (das kann auf eine obere Altersgrenze zum Screening hinauslaufen, die mit Altersrationierung von Leistungen nichts zu tun hat).

Das Angebot zur Teilnahme an einem positiv vorbewerteten Früherkennungsprogramm kann mehr oder weniger nachdrücklich vermittelt werden, von einer kaum kommunizierten „Anspruchsberechtigung“ bis hin zu systematischen Einladungsmodellen, wie sie in Deutschland für das mammografische Screening auf Brustkrebs eingerichtet und für das Darmkrebs- und Gebärmutterhals-Screening beschlossen wurden. Schon die Erfahrungen mit dem Gebärmutterhals-Screening haben gezeigt, dass es wirksamer ist, viele Anspruchsberechtigte mindestens einmal zu screenen als wenige repetitiv. Dies spricht dafür, nicht allein auf die Teilnahme immer derselben selbstmotivierten Freiwilligen zu setzen, sondern nachdrücklich für breite Inanspruchnahme zu werben, gerade auch in schwer erreichbaren Risikogruppen. Allerdings befinden sich die Gestaltungsverantwortlichen dabei in einem Dilemma. Je nachdrücklicher die Werbung, desto mehr wird die Selbstbestimmung der Versicherten durch Paternalismus übersteuert, dem man zwar die gute Absicht nicht absprechen muss, der aber wie oben gezeigt der individuellen Aufklärungsnotwendigkeit nicht gerecht wird und individuelle Präferenz nicht mehr respektiert.

Auch eine hundertprozentige Beteiligung am Screening kann die Gesamtsterblichkeit nach einer Modellrechnung nur gering senken, beim mammografischen Screening um 1,8%, bei den

² In § 25 Abs. 1 SGB V heißt es „bevölkerungsmedizinisch bedeutsame Krankheiten“.

anderen Krebslokalisationen deutlich weniger (Stang u. Jöckel 2018). Deswegen und wegen der Verschlechterung der Wirtschaftlichkeit durch ein ungünstiges Verhältnis von Vorhaltekosten zu gesundheitlichen Effekten diskutiert man eine Teilnahme-Untergrenze, ab der ein Screening-Programm für Bevölkerung und Gesundheitswesen sinnvoll wird. Bei Einführung des mammografischen Screenings war deswegen unter Bezug auf eine EU-Leitlinie eine (zweijährige) Teilnahmerate von mindestens 70% als Randbedingung für die Ein- und Fortführung des Programms angestrebt worden. Im letztverfügbaren Jahresbericht der Kooperationsgemeinschaft Mammografie lag die Einladungsrate der Anspruchsberechtigten 2015 bei praktisch 100%, die Teilnahmerate bei knapp 52% (überdurchschnittlich in den neuen Bundesländern) (Kooperationsgemeinschaft Mammografie 2017).

Die Politik hat nach breiter (fach-)öffentlicher Debatte in dieser Hinsicht einen Paradigmenwechsel vollzogen (s. Kap. 6). Vorrang vor der Verbesserung der Teilnahmeraten als intermediärer Zielgröße des Programms hat die verbesserte Information der Bürgerinnen und Bürger über Nutzen und Risiken der Krebsfrüherkennung mit dem Ziel einer informierten Entscheidung. Damit ist das bislang weithin konsentiierte „Werben“ für die Teilnahme im Kern obsolet. Wird eine informierte Entscheidung im Alltag gelebt, können die Teilnahmeraten sogar sinken, was unbedingt zu respektieren wäre. Es steht zu vermuten, dass dieser Wandel noch längere Zeiträume in Anspruch nehmen wird.

Ein geplantes Screening-Programm gibt Teilnahmeintervalle vor (siehe z.B. die komplexen Vorgaben zum Darmkrebs-Screening in Teil II, § 3 der o.g. oKFE-RL). Allgemein gesagt ist die Wiederholungsfrequenz des Screenings abhängig von der Verteilung der Wachstumsgeschwindigkeiten der Zielläsionen und der methodenabhängigen (durchschnittlichen) Vorverlegung des Diagnosezeitpunkts in der Früherkennung („lead time“, Hill et al. 1984). Das in Deutschland zu Beginn der organisierten Früherkennung für alle Zielläsionen und Maßnah-

men einheitlich festgelegte jährliche Intervall ist offensichtlich zu schematisch und wird für jüngere Screening-Verfahren wie die Mammografie und die Koloskopie, zukünftig auch für das Screening auf Gebärmutterhalskrebs und den Check-up 35, ausdifferenziert. Internationale Empfehlungen halten seit Jahren beim Cervix-Screening ein 3- bis 5-jähriges Intervall für ausreichend, wenn die Qualität der Abstriche gut ist; bei mehrfach negativ gescreenten Teilnehmerinnen kann auf weitere Untersuchungen verzichtet werden (Hakama et al. 1986; Curry et al. 2018). Die Überarbeitung des Früherkennungsprogramms auf Cervixkarzinome hinsichtlich Altersgrenzen und Intervallen wird zudem wegen der Möglichkeit, HPV-Infektionen nach Risiko zu differenzieren, und der zusätzlich verfügbaren, gerade auf Jungen ausgetesteten HPV-Impfung dringlich. Ein Konzept des G-BA für eine neu organisierte Früherkennung auf Gebärmutterhalskrebs³ und ein Evaluationskonzept des IQTiG (2017) liegen vor.

Zu c) die aufgeklärte Entscheidung einer symptomlosen Person zur Teilnahme

Diese Entscheidung können Versichertengemeinschaft und Individuum nur treffen, wenn gute wissenschaftliche Evidenz zu Organisation, Mengengerüsten, richtig positiven und falsch positiven Folgemaßnahmen und positiven wie negativen gesundheitlichen Langzeiteffekten erarbeitet wurde und mit Hinweisen auf ihre Aussagefähigkeit verständlich aufbereitet zur Verfügung steht. Die methodischen Anforderungen an vertrauenswürdige Screening-Programme sind also hoch. Allgemein sind die Qualität von Informationsmaterialien und Aufklärungsgesprächen und das Wissen der Bürger jedoch unzureichend (Dierks u. Schmacke 2014; Seidel et al. 2014; s. Kap. 5). Da

3 Pressemitteilung vom 14. April 2018: https://www.g-ba.de/downloads/34-215-742/14_2018-04-17_KFE-RL_Zervixkarzinom-Screening.pdf

jedes neue diagnostische oder therapeutische Verfahren die Risiko-Nutzen-Balance der Früherkennung verschiebt, müssen Informationen periodisch aktualisiert werden.⁴

Wenn ein Individuum wie gesehen für seine Entscheidung, eine Früherkennungsmaßnahme in Anspruch zu nehmen, vor dem Schleier der Ungewissheit über die für ihn folgende Ereigniskette gerade dieselbe quantitative Evidenz wie die Versichertengemeinschaft als Ganze braucht, jedoch die nachfolgenden Maßnahmen und gesundheitlichen Resultate anders bewerten kann als die Versichertengemeinschaft, braucht es zusätzlich Unterstützung zur Klärung seiner individuellen Präferenzen, jedenfalls solange es Gründe sehen kann, sich nicht dem gesellschaftlichen Hauptstrom anzuschließen.

Ein solcher Grund könnte z.B. eine einschlägige genetische Belastung in der Familie sein, die oft mit entsprechenden Krankheitserfahrungen bei Angehörigen verbunden ist. Personen mit solchen Zusatzrisiken können präferenzbedingt leichter zu Früherkennungsmaßnahmen und intensiveren Interventionen motivierbar sein als Versicherte mit durchschnittlichem Risiko. Ob sie daraus allerdings mehr Nutzen realisieren können als eine Person durchschnittlichen Risikos, ist eine andere, gesondert zu untersuchende Frage.

Zu d) die Durchführung notwendiger Folgeuntersuchungen zur Abklärung eines auffälligen Früherkennungstests

Auf welche Weise auffällige Screening-Tests abgeklärt werden, ist untrennbarer Bestandteil jeder Früherkennungsstrategie. Ob z.B. die Spiegeleinstellung des Gebärmuttermundes schon als Teil des zytologischen Screenings unter Sicht

eines Mikroskops (Kolposkop) durchgeführt wird (Beller et al. 1982) oder ob eine Kolposkopie erst nach zytologischer Auffälligkeit zur Kontrolluntersuchung gehört, führt zu unterschiedlichen Ergebnissen. Bei welchen Befunden im Brustkrebs-Screening zusätzlich eine Ultraschalluntersuchung angezeigt ist, ist Gegenstand von Kosten-Effektivitäts-Analysen (z.B. Sprague et al. 2015). Mit welcher Technik der Prostata nach einem auffälligen PSA-Test Gewebe entnommen wird, hat Auswirkungen auf Entdeckungsraten (s. Kap. 10). Mit welcher Verzögerung eine empfohlene Biopsie durchgeführt wird, ist Parameter der Qualitätssicherung der Früherkennung (Meyer et al. 1997). Systematisch eingeholte Zweitmeinungen für auffällige Pathologiebefunde können die Behandlungsindikation präzisieren und sogar Kosten sparen (Tosteson et al. 2018). Da ständig diagnostische Fortschritte möglich werden, ist der Diskurs um die beste (wirksamste, wirtschaftlichste) Abklärungsstrategie im Screening grundsätzlich offen. Fortschritte der Abklärungsstrategie müssen zu einer Neubewertung vorangehender Schritte der Früherkennung führen.

Da in der GKV Abklärungsmaßnahmen normalerweise „kurativ“ abgerechnet werden, ist bei einer evaluativen Auswertung von Abrechnungsstatistiken nicht immer klar, ob eine Maßnahme Folge einer Früherkennungsuntersuchung oder einer Abklärung von Symptomen gewesen ist. Der Bewertungsausschuss hat daher für Früherkennungsleistungen je spezielle Gebührenordnungspositionen im Einheitlichen Bewertungsmaßstab festgelegt (z.B. GOP 01733 für zytologische Untersuchung nach KFE-RL, 01741 für Koloskopie nach KFE-RL oder 01748 für Sonografie nach US-BAA-RL). Krebsregister bemühen sich um eine genaue Zuordnung der Maßnahmen.

Zu e) die indizierte Frühintervention

Seit langem ist klar, dass Erfolgsstatistiken der klinisch-kurativen Individualmedizin für die

⁴ Neben den Krankenkassen hat das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) den Auftrag, über Vorteile und Nachteile von Untersuchungs- und Behandlungsverfahren zu informieren (www.iqwig.de). Auch die DEGAM stellt Entscheidungshilfen zur Verfügung (DEGAM 2018; Kötter 2016).

Bewertung von Früherkennungsprogrammen nicht gültig sind. Dies trifft besonders auf die 5-Jahres-Überlebensrate zu (McNeil et al. 1978), aber auch auf eine Vorverlegung der Stadienverteilung. Wesentliche Gründe dafür liegen in der (intendierten) Vorverlegung des Diagnosezeitpunkts gegenüber der üblichen Versorgung selber („lead time“) und in der Tendenz, in Filteruntersuchungen überproportional langsam wachsende, damit ohnehin prognostisch günstige Veränderungen zu entdecken („length bias“). Hinzu kommt eine Selbstselektion bei der Inanspruchnahme („healthy screenee bias“), die den Vergleich der Ergebnisse von Früherkennungsteilnehmern und Nichtteilnehmern verzerrt.

Für die zentrale Frage, ob Früherkennung die Prognose ihrer Zielläsion verbessert, sind daher nicht klinische, sondern bevölkerungsbezogene Studien und deren Auswertung auf Basis der zur Früherkennung eingeladenen Bevölkerung („intention to screen“) erforderlich, so das Monitum von Epidemiologen und Versorgungsforschern. Die erste derartige Studie, eine regional-flächendeckende Studie zum Mammografie-Screening in Schweden ab 1977, publizierte erste Ergebnisse 1985 (Tabár et al. 1985). Die Studie bestätigte das Früherkennungsprinzip beim Brustkrebs und gilt als einer der Ecksteine der Brustkrebsfrüherkennung. Allerdings wurde sie nach Überprüfung zeitgenössischer schwedischer Krebsregisterdaten als fehlerhaft eingestuft (Zahl et al. 2006).

Inzwischen sind weitere kontrollierte Studien zum Screening verschiedener Zielläsionen erschienen, darunter solche, deren Durchführung bemängelt wird. Die leidenschaftlichste Kontroverse wird über die Grundlagen des Mammografie-Screenings geführt. Quanstrum und Hayword sprechen von „mammography wars“ (2010). Zahlreiche unterschiedliche „Stakeholder“ wie Patientengruppen, Laborwissenschaftler, Kliniker, Manager und Politiker bringen Interessenpositionen ein (Armstrong u. Eborall 2012). Die Leidenschaft der Debatten

verweist auf ihre Kerndimensionen Tradition und Macht (nicht: Evidenz).

Auch für Experten ist die Belastbarkeit aller Argumente pro und contra kaum noch überschaubar. Zudem ist die Übertragbarkeit der Ergebnisse lange zurückliegender Studien aus anderen Gesundheitssystemen auf die aktuelle Versorgung in der deutschen Krankenversicherung eingeschränkt. Die Film-Mammografie der Tabár-Studie, die damaligen Abklärungsverfahren und Therapieschemata sind z.B. inzwischen überholt. Wir müssen daher ernüchtert feststellen, dass auch bevölkerungsbezogene Versorgungsforschung und der Versuch methodisch stringenter Evidenzbasierung fehlbar sind. Das wiederum kann nicht bedeuten, einseitigen politisch motivierten oder interessengeleiteten Positionen Raum zu geben, sondern muss zu noch besser reflektierter und kontrollierter Versorgungsforschung Anlass geben (siehe unten).

Beim Mammografie-Screening allerdings griff die Politik in den meisten Ländern die Argumente der Befürworter auf: In sich immer wiederholenden Statements wird mit einer 30%-igen Senkung der Sterblichkeit (anstelle des absoluten Risikos) argumentiert (siehe z.B. die fraktionsübergreifende Beschlussempfehlung aus dem Gesundheitsausschuss des deutschen Bundestages zur Einführung eines solchen Screenings, Deutscher Bundestag 2002). Zwar standen den starken Pro-Mammografie-Kampagnen im Politikfeld immer auch kritische Stimmen entgegen, die vor einer Pathologisierung weiblicher Lebensphasen warnen (z.B. Tormann 2013). Sie konnten sich im gesundheitspolitischen Raum aber nicht behaupten.

Zu f) die begleitende Dokumentation, Erfolgssicherung und Bewertung (Evaluation) des Vorgehens

Es ist offensichtlich unzulässig, ein Programm, d.h. einen Plan oder ein „Soll“, mit seiner tat-

sächlichen Durchführung, dem „Ist“, gleichzusetzen. Allgemein hat in der Medizin – trotz Sorgfalt in Vorbereitung und Durchführung – das erfahrungsgestützte Lernen am Erfolg einen hohen Stellenwert. Die Evaluation eines Früherkennungsprogramms ex ante ist oben beschrieben. Wie kann man aber ein schon laufendes Früherkennungsprogramm (ex post) bewerten? Läuft es wie vorgesehen, erreicht es (vorläufig) seine Ziele, sind korrigierende Eingriffe nötig?

Das zytologische Screening auf Cervixkarzinome und seine Vorstufen mit dem „Pap-Test“ gilt als Erfolgs-Prototyp des Krebs-Screenings. Sein positiver Beitrag zum Rückgang des Gebärmutterhalskrebses ist anerkannt, auch wenn es nie randomisierte Vergleiche gegeben hat. Trotzdem gibt es Kritik an der Wirksamkeit und Wirtschaftlichkeit der Umsetzung (z.B. Charny et al. 1987; Murphy et al. 1988). Die nicht ausevaluierte Plausibilität dieses Tests lässt extreme Unterschiede in der Dichte seiner Durchführung zu, auch wenn gezeigt werden konnte, dass die Intensivierung des Screenings in keinem positiven Verhältnis zum Outcome steht (siehe z.B. Habbema et al. 2012).

Da klinische Verlaufsdokumentationen wie dargelegt zur Erfolgsbewertung von Früherkennungsprogrammen verzerrt sind und nur Indizien beitragen können, stützt sich die laufende Evaluation wesentlich auf die Routinedokumentation zur Qualitätssicherung der Programmdurchführung, auf die sekundäre Auswertung von Abrechnungsdaten sowie auf Register- und Mortalitätsdaten. Für das Mammografie-Screening haben KBV und GKV-Spitzenverbände dafür die „Kooperationsgemeinschaft Mammografie“ gegründet (www.mammo-programm.de). Das Krebsregister- und Früherkennungsgesetz (s. Kap. 6) verfolgt das Ziel, die Datengrundlagen zur Zusammenführung verteilter Informationen zu verbessern. Der von Brunßen et al. (s. Kap. 13) dargelegte Stand zur laufenden Evaluation des Hautkrebs-Screenings lässt erkennen, mit welcher Umsicht solche Daten interpretiert werden müssen.

Das jüngst vom G-BA beschlossene Screening auf Bauchaortenaneurysma bei Männern ab 65 Jahren stützt sich auf eine vorbereitende Ex-ante-Evaluation durch das IQWiG und den Unterausschuss Methodenbewertung (Gemeinsamer Bundesausschuss 2016). Sie weicht von der vorausgehenden Empfehlung der United States Preventive Services Task Force (USPSTF; LeFevre 2014) durch Fehlen einer oberen Altersgrenze ab. Zusätzlich hat der G-BA eine evidenzbasierte Entscheidungshilfe für die Versicherten (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen 2016) und eine unabhängige wissenschaftliche Evaluation nach drei Jahren vorgesehen. Die Evaluation soll Teilnahmeraten, Krankenhausdiagnose- sowie Todesursachenstatistik, die Mengenentwicklung elektiver Operationen und Notfalloperationen sowie die Entwicklung der Aneurysma-assoziierten Mortalität betrachten (Bundesministerium für Gesundheit 2017). Ein solcher externer Evaluationsauftrag ist ein vernünftiger Schritt im Interesse einer formativen „Politik ohne Reue“, ersetzt aber nicht eine belastbare Evaluation ex ante.

1.3 Gemeinsame Entscheidungsfindung, Health Literacy und Respekt vor dem Bürger

Entscheidungen zum medizinischen Handeln sind nicht immer eindeutig indiziert, sondern es gibt zahlreiche Grauzonen des jetzt Möglichen, hier Verfügbaren, wahrscheinlich Wirkenden, für diesen Patienten hinsichtlich Ergebnis und unerwünschter Wirkungen (noch) Akzeptablen oder (schon) Abgelehnten und auch für den einzelnen Arzt Verantwortbaren (Robra u. Spura 2018). Diese Individualität der Versorgung ist grundsätzlich erwünscht.

Seit gut zwanzig Jahren wird in den Gesundheitswissenschaften diskutiert, warum es so schwer ist, die ‚Zielgruppen‘ des medizinischen Screenings zu bemächtigen, informiert selber darüber zu entscheiden, ob sie angebotene

Screening-Programme nutzen wollen oder nicht. In einer frühen Veröffentlichung hierzu hat Austoker (1999) nicht nur darauf verwiesen, dass Screening zwangsläufig auch mit – in der Regel massiv unterschätzten – körperlichen und seelischen Nebenwirkungen verbunden ist, sondern dass vor allem Instrumente fehlen, die ‚Zielgruppen‘ über diese komplexen Zusammenhänge auf ihrem jeweiligen Verständnisniveau so aufzuklären, dass stabile Haltungen pro oder contra entstehen können. Die unstrittige Tatsache, dass bevölkerungsbezogene auf den Gesundheitszustand abzielende, öffentlich unterstützte Interventionen dennoch mit einem Schadenspotenzial verbunden sind, das in der Gegenüberstellung zum Nutzen vertretbar sein muss, ist offenkundig für methodisch nicht geschulte Menschen in hohem Maße kontraintuitiv. Individualpsychologisch spielt außerdem das Nachdenken über mögliches Selbstverschulden von Krankheit eine Rolle. Nichtteilnahme an einem Screening kann dann als Unterlassen nötiger Eigenverantwortung gedeutet werden.

Wiederholt konnte empirisch gezeigt werden, dass der Nutzen des Krebs-Screenings massiv über- und das Schadenspotenzial massiv unterschätzt wird. Dierks und Schmacke (2014) deuteten eine Befragung von GKV-versicherten Frauen aus dem Jahr 2014 wie folgt:

„Jede zweite Frau ist falsch oder unzureichend informiert, wenn es um Früherkennung von Brustkrebs oder Mammografie-Screening geht. [...] So glauben 30 Prozent der Frauen, dass die bloße Teilnahme am Mammografie-Screening verhindert, dass sie an Brustkrebs erkranken. Der Nutzen der Untersuchung wird überschätzt, während über die Risiken etwa durch falsche Positiv-Befunde nur wenig bekannt ist.“

Es war diesbezüglich kein Fortschritt gegenüber der europäischen Vergleichsstudie von Gigerenzer et al. zu verzeichnen (Gigerenzer et al. 2009). Auf das generell massive Missverhältnis von Evidenz und positiver Einschätzung präventivmedizinisch motivierter Unter-

suchungen hat Mühlhauser 2014 in einer Übersichtsarbeit hingewiesen (siehe auch Sommer et al. 2018). Dies ist ein treibender Faktor für die Existenz von ‚grauem‘ Screening in allen Gesundheitssystemen: Die Ärzteschaft bietet immer auch unsystematisch Untersuchungen unter dem Versprechen der Verhütung von sonst zu spät erkannten Erkrankungen an und kann dabei auf diffuse Ängste in relevanten Teilen der Bevölkerung bauen.

Noch immer geht es um basale Fragen: Wer etwa denkt bis ans Ende der Screening-Kette? Der PSA-Basiswert kommt als unproblematischer Bluttest daher, ein minimaler Eingriff. Dass danach die Aufforderung zu einer oder mehrfachen Punktionen der Vorsteherdrüse folgen kann, ist gedanklich weit weg. Und völlig unvorbereitet kann die Männer dann die Entscheidung treffen: Abwarten oder Behandeln, wenn schwarz auf weiß von Krebszellen die Rede ist. Natürlich müsste Aufklärung schon vor der Blutentnahme solche Ketten reflektieren. Denn: Im günstigsten Fall geht es „nur“ um den eingebauten Passagenwechsel vom Gesunden zum Überwachten. Im übelsten Falle um unnötige Operationen oder Bestrahlungen oder medikamentöse Behandlung mit all ihren Risiken. Und nach erfolgter Operation oder Bestrahlung zählt wohl oft nur der von verständlichen Hoffnungen getragene Gedanke: Gut, dass der Krebs weg ist. Die Idee, mit einem diagnostizierten Krebs zu leben und ihn „nur“ überwachen zu lassen, ist aus wissenschaftlicher Sicht für viele Männer eine denkbar gute Idee. Gefühlt kann sie aber kaum erträglich sein. Wie bewerkstelligt werden kann, dass diese Gedankenkette tatsächlich Teil des Redens über Prostatakrebs wird, ist momentan wohl noch weitgehend unklar.

Eine Kernfrage für eine wünschenswerte, wissenschaftlichen und ethischen Maßstäben gerecht werdende Weiterentwicklung von Screening-Programmen lautet insofern: Wie bekommen Menschen eine Vorstellung über die Dimensionen von Nutzen und Schaden? Vereinfacht gesagt: Es gibt zwar erste methodisch an-

spruchsvoll erstellte Entscheidungshilfen für Ratsuchende bzw. Anspruchsberechtigte, es ist aber weitgehend unklar, wie daraus eine neue Kultur der selbstbestimmten Entscheidung erwachsen kann. Weder der gewohnte Paternalismus in der Arzt-Patient-Beziehung noch das tief verankerte Denken „früh kann nur gut sein“ und die daran gekoppelten Ängste vor dem Verpassen bedeutender Gesundheitschancen sind in kurzen Zeiträumen in eine neue Beziehung auf Augenhöhe zwischen Experten und Laien und eine Kompetenz zum selbstsicheren Betreten eigener Wege mit dem oder gegen den Strom zu transformieren.

Was sollte geschehen, um im nächsten Jahrzehnt (oder in einem noch längeren Zeitraum?) ein solideres Fundament für die Beratung zu Screening und die individuelle Entscheidung pro oder contra zu schaffen? Wer sollte dabei welche Rolle spielen, wenn offenkundig ist, dass das Thema weit über den medizinischen Alltag hinausweist?

Die informierte Patientenentscheidung und die partizipative Entscheidungsfindung gelten seit geraumer Zeit als die Schlüssel zum Durchbruch rationaler wie patientenzentrierter Medizin (s. Kap. 5). Alles rankt sich um den Begriff der ‚Gesundheitskompetenz‘ (health literacy). Es geht, wie Bitzer und Sørensen (2018) in einem lesenswerten Fortbildungstext gesagt haben, um

„die Fähigkeit, gesundheitsrelevante Informationen aus verschiedenen Quellen zu finden, zu bewerten und für eigene gesundheitsbezogene Entscheidungen anzuwenden sowie die Fähigkeit, selbst gesundheitsrelevante Informationen zu kommunizieren, bspw. in der Patient-Arzt-Kommunikation“.

Mit anderen Worten: Es geht zum einen um einen fairen Zugang zu den jeweils relevanten Gesundheitsinformationen, denen Vertrauen geschenkt werden darf. Und es geht zum anderen um die gesellschaftliche Einübung sich darauf stützender autonomer Entscheidungen in einem letztlich dauerhaft asymmetrischen Ver-

hältnis zwischen Akteuren und Nutzern des Versorgungssystems.

Es gibt in diesem Zusammenhang eine Reihe von weitgehend unangetasteten Vorannahmen, deren Gültigkeit aber infrage steht. Dazu gehört die Denkfigur „Experten versus Laien“. So unstrittig es ist, dass Ärztinnen und Ärzte in der Regel einen großen Wissensvorsprung vor ihren Patientinnen und Patienten haben (es sei beiseite gelassen, dass diese Unterschiede im Falle chronischer Krankheit nicht selten schmelzen), so sehr wird verkannt, dass die Aus-, Fort- und Weiterbildung der Ärzteschaft vor dem Problem steht, wissenschaftliche Grundsätze für das Alltagshandeln fruchtbar zu machen: Hierzu zählen Fragen der Wahrscheinlichkeitstheorie und die ‚Verurteilung‘ zum Handeln unter Ungewissheit (s. Scheff 1963; Wegwarth 2013; Makridakis et al. 2018). In einem nicht konstruierten Extremfall können so ein gut informierter Patient und ein schlecht informierter Arzt über das PSA-Screening sprechen, womit das Prinzip für eine gemeinsame Entscheidungsfindung ad absurdum geführt wird. Im anderen Extremfall treffen ein exzellent informierter Arzt und ein gesundheitlich illiterater Patient aufeinander – und es fragt sich, ob für diese Situation evidenzbasierte und für das Verständnisniveau des Patienten geeignete Entscheidungshilfen zur Verfügung stehen. Die Entwicklung von Decision Aids für Screening, Diagnostik und Behandlung kann prinzipiell als großer Hoffnungsträger gelten (Stacey et al. 2017), der Aufgabenkatalog einschließlich der Berücksichtigung unterschiedlicher Bildungsvoraussetzungen ist freilich groß. In Deutschland wurden Anforderungen an die methodischen Standards wohl erstmals 2005 von der Arbeitsgruppe von Mühlhauser begründet publiziert (Steckelberg et al. 2005). Es besteht weitgehend Konsens über die stufenweise Entwicklung von der Literaturrecherche über die Einbindung von Laien bis hin zum Testen des Piloten im Feld. Ein Überblick über die Begrifflichkeiten und das breite Feld von Anbietern unterschiedlicher

Formate von Patienteninformationen findet sich bei Lenz et al. (2012). Zur Aufklärung von Laien wie Experten hat die Verwendung natürlicher Zahlen (im Gegensatz zu Wahrscheinlichkeiten) eine herausgehobene Bedeutung – inzwischen ist dies Standard evidenzbasierter Entscheidungshilfen (Lühnen et al. 2017).

Weiter zu verfolgen ist die Frage, welchen Einfluss evidenzbasierte Entscheidungshilfen im Versorgungsalltag haben, inwieweit sie vor allem das Ziel erreichen, harte klinische Ergebnisse und die Patientenzufriedenheit zu verbessern. Der Zeitraum, hierzu verlässliche Aussagen für eine Vielzahl von Erkrankungen – inkl. Screening-Programmen – machen zu können, ist noch zu kurz. Die Cochrane Collaboration hat hierzu im Lichte der vorhandenen Evidenz ein Studienprotokoll entwickelt (Perrenoud et al. 2015). Eine einfache Informationsbroschüre ist jedenfalls nicht ausreichend (Gummersbach et al. 2015).

Allerdings sind die Hürden für eine angemessen wirkende Kommunikation in der Arzt-Patient-Beziehung nach wie vor sehr hoch (Schmacke et al. 2016). Im Falle des Krebs-Screenings ist auch der Fortbildungsbedarf für Ärztinnen und Ärzte wiederholt empirisch nachgewiesen worden (z.B. Frank et al. 2018; Wegwarth u. Gigerenzer 2018). Jenseits ambitionierter Debatten etwa um Kausalitätsnachweise oder ethische Anforderungen an Screening ‚lebt‘ der Screening-Diskurs von Narration: „Aber ich habe gerade in meiner Praxis gesehen ...“ oder: „Ich habe gerade im Freundinnenkreis gehört, dass ...“. Dahinter verschwinden die häufig parallel bestehenden diffusen Zweifel, ob es wirklich nützt.

Fortbildungsbedarf gilt nicht zuletzt für neuere Gentests (Chromosomal microarray analysis CMA) in der Kinderheilkunde (Reiff et al. 2013). Die rasante Entwicklung des Neugeborenen-Screenings (s. Kap. 7) verweist auch noch einmal auf die ethische Dimension von Screening: Timmermans und Buchbinder sprechen angesichts der hohen prognostischen Unsicherheiten vieler dieser Tests von ‚Patients-in-Waiting‘:

Kinder schweben zwischen den Prognosen Gesundheit und Krankheit:

“The social significance of diagnosis has always been its functionality: the ability to structure medical encounters, institute health policies, shape life strategies, even control deviance. In the case of patients-in-waiting, this social significance has now been extended to those living between health and disease” (Timmermans u. Buchbinder 2010, 420).

Die Ära der Genomics und Proteomics hat längst ein neues, in seinen Folgen noch weiterhin unverstandenes Kapitel der Suche nach Assoziationen zwischen Markern und Krankheitsrisiken aufgeschlagen.

Eine weitere bedeutende Facette im Kontext ‚autonome Entscheidung‘ bezieht sich auf die Form der Einladung zum Screening (siehe oben), wobei Trevena (2009) darauf hinweist, dass in jedem Fall ausreichend Zeit vorhanden sein muss, nach einer Basisinformation über Pro und Contra weiterführende Informationen einzuholen, wenn dies gewünscht ist: ein nach wie vor unterschätztes Argument für Opt-In-Lösungen, die momentan gegenüber dem simultanen Versenden von Einladungen und Aufklärungsmaterialien noch kaum diskutiert werden. Vermutlich spielen nach wie vor Grundhaltungen eines ‚benevolenten Paternalismus‘ gegenüber der Idee der Förderung autonomer Entscheidungen gerade in der Ärzteschaft eine dominierende Rolle (Jørgensen 2013). Die Rede ist zunehmend von einer gewissermaßen liebevollen Form des Überredens zum Screening, dem Nudging (Ploug et al. 2012) – aus ethischer Sicht ebenso wenig akzeptabel wie frühere angstgesteuerte Aufforderungen zur Teilnahme (eine gegenteilige Auffassung aus Sicht der Bioethik vertritt Cohen 2013).

Die deutsche Gesundheitspolitik hatte demgegenüber auf die Teilnahme an Früherkennungsuntersuchungen direkten Druck ausüben wollen, indem im Rahmen der Zuzahlungs- und Überforderungsregelungen bei Nichtteilnahme die Zuzahlungsgrenze für chronisch Kranke die

Höchstbeteiligung von 2% vom Bruttoeinkommen gelten sollte (§ 62 SGB V). Der von der Politik eingeschaltete Gemeinsame Bundesausschuss hat dies 2004 im Grundsatz nicht moniert, sondern nach einer patientengerechten, aber auch für die Politik gesichtswahrenden Lösung gesucht, die darin besteht, dass die betroffenen Versichertengruppen sich über „Chancen und Risiken“ der Früherkennung ärztlich beraten lassen müssen, um von der abgesenkten 1%-Regelung profitieren zu können (hierzu umfassend aus ethischer, sozialrechtlicher und ökonomischer Sicht Schmidt 2012). Die Denkmuster in der Politik waren jedenfalls darauf aus, mit einer Malus-Regelung die Teilnahme vor allem an den Krebsfrüherkennungsuntersuchungen zu erhöhen. Erst mit dem Krebsfrüherkennungs- und -registergesetz wurde erstmals in der deutschen Gesundheitspolitik der Vorrang der informierten Entscheidung verankert:

„Ein zentrales Anliegen ist die Stärkung der ‚informierten Entscheidung‘ der gesetzlich Versicherten mithilfe fundierter und ausgewogener Informationen über die Vor- und Nachteile der jeweiligen Krebsfrüherkennungsmaßnahme. Die Priorisierung der informierten Entscheidung über die reine Teilnahmesteigerung kann als Paradigmenwechsel in der deutschen Gesundheitspolitik betrachtet werden“ (Helou 2014).

Die Intention des ausdrücklichen Werbens für das Krebs-Screening ist dem Gesetz gleichwohl weiter anzumerken, soll dieses doch Schritt für Schritt auf ein Einladungsverfahren umgestellt werden: Das Gesetz ist insoweit janusköpfig. Auch ein Bonus für erfolgte Krebsfrüherkennungsuntersuchungen ist kritisch zu sehen.

Dies alles hat Folgen für die medizinische Ausbildung, Weiterbildung und die Organisation des Gesundheitswesens. Der Prozess der Zielklärung auf der Individualebene ist zu selten Gegenstand der überwiegend krankheitsbezogenen medizinischen Ausbildung. Das soll sich nach dem „Masterplan2020“ ändern. Man wird aber davon ausgehen müssen, dass in Poli-

tik, Gesellschaft allgemein und im Gesundheitswesen im speziellen alte Denkmuster in Sachen Screening noch längere Zeit wirkmächtig bleiben.

Der G-BA als zentraler, politisch autorisierter Akteur der Umsetzung der politischen Screening- und Präventionsstrategien steht trotz seiner evidenzorientierten Verfahrensordnung (Gemeinsamer Bundesausschuss 2018) fraglos unter starkem Druck. Zum einen durch Partialinteressen, wenn er sich auf bestverfügbare Evidenz stützt und sich nicht selten dem Vorwurf ausgesetzt sieht, zu restriktiv zu entscheiden. Zum zweiten sind mit der gesetzlich fixierten Zusammensetzung des Entscheidungsgremiums unabweisbar Interessenkonflikte institutionalisiert. Drittens hat der G-BA politische Vorentscheidungen zu respektieren, auch wenn er sie inhaltlich nicht oder nicht vollständig mitträgt. Andererseits sollen seine Entscheidungen den Leistungskatalog der gesetzlichen Krankenversicherung evidenzbasiert weiterentwickeln, das Gesundheitswesen wirksamer und wirtschaftlicher organisieren und die Gesundheitskompetenz in der Gesellschaft fördern. Welche fachlichen Entscheidungsvorbereitungen er anfordert und erhält und wie er Politik und Gesellschaft über seine Entscheidungen, ihre tragenden Gründe und Limitationen informiert, trägt dazu bei.

Literatur

- Armaroli P, Villain P, Suonio E, Almonte M, Anttila A, Atkin WS et al. European Code against Cancer, 4th Edition: Cancer screening. *Cancer Epidemiol* 2015; 39 Suppl 1: S139–52.
- Armstrong N, Eborall H. The sociology of medical screening: past, present and future. *Sociology of Health & Illness* 2012; 34:161–176.
- Austoker J. Gaining informed consent for screening is difficult – but many misconceptions need to be undone. *BMJ*. 1999; 319:722–723.
- Autier P, Boniol M, Koehlin A, Pizot C, Boniol M. Effectiveness of and overdiagnosis from mammography screening in the Netherlands: population based study. *BMJ* 2017; 359:j5224.
- AWMF Leitlinienprogramm Onkologie. S3-Leitlinie Prostatakarzinom, Version 5.0, April 2018; 2018. Available from: URL: <https://www.leitlinienprogramm-onkologie.de/leitlinien/prostatakarzinom/>.

- Beller FK, Brecht JG, Schmidt E-H, Schwartz FW. Nutzen der Kolposkopie bei Krebsfrüherkennungs – Untersuchungen: Westfalen-Lippe-Studie. *Deutsches Ärzteblatt* 1982; 79(9):A33-A38. doi: 10.1001/jamainternmed.2018.2773.
- Bhatia S, Dorian P. Screening for Cardiovascular Disease Risk With Electrocardiography. *JAMA Intern Med.* 2018; 178(9):1163-1164. doi: 10.1001/jamainternmed.2018.2773.
- Bill-Axelsson A, Holmberg L, Garmo H, Rider JR, Taari K, Busch C et al. Radical prostatectomy or watchful waiting in early prostate cancer. *N Engl J Med* 2014; 370(10):932-42.
- Bitzer EM, Sørensen K. Gesundheitskompetenz – Health Literacy. *Gesundheitswesen* 2018; 80(8-09):754-66.
- Bundesministerium für Gesundheit. Bekanntmachung eines Beschlusses des Gemeinsamen Bundesausschusses über eine Änderung der Krebsfrüherkennungs-Richtlinie: Bewertung eines iFOBT-basierten Darmkrebscreenings im Vergleich zu einem gFOBT-basierten Darmkrebscreening vom: 21.04.2016; BANz AT 08.07.2016 B2.
- Bundesministerium für Gesundheit. Bekanntmachung eines Beschlusses des Gemeinsamen Bundesausschusses über eine Richtlinie Ultraschallscreening auf Bauchaortenaneurysmen vom: 20.10.2016; BANz AT 17.02.2017 B4.
- Caley M, Chohan P, Hooper J, Wright N. The impact of NHS Health Checks on the prevalence of disease in general practices: a controlled study. *Br J Gen Pract* 2014; 64:e516-21.
- Cancer Council Australia's National Skin Cancer Committee. Position statement – Screening and early detection of skin cancer: This position statement is endorsed by the Australasian College of Dermatologists; July 2014. https://wiki.cancer.org.au/policy_mw/index.php?oldid=6289 (30.07.2018)
- Charny MC, Farrow SC, Roberts CJ. The cost of saving a life through cervical cytology screening: implications for health policy. *Health Policy* 1987; 7(3):345-59.
- Chenot J-F, Egidi G. Empfehlungen zum Hautkrebscreening in der S3-Leitlinie „Prävention von Hautkrebs“ – Kritik der DEGAM Teil 2. *Z Allg Med* 2015; 91(3):121-5.
- Cohen S. Nudging and informed consent. *Am J Bioeth* 2013; 13:3-11.
- Correa C, McGale P, Taylor C, Wang Y, Clarke M, Davies C et al. Overview of the randomized trials of radiotherapy in ductal carcinoma in situ of the breast. *J Natl Cancer Inst Monogr* 2010; 2010(41):162-77.
- Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, Davidson KW et al. Screening for Cervical Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA* 2018; 320(7):674-86.
- Davis TC, Williams MV, Marin E, Parker RM, Glass J. Health Literacy and Cancer Communication. *CA* 2002; 52:134-149.
- DEGAM Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin e.V. Hausärztliche Beratung zu PSA-Screening: DEGAM-Praxisempfehlung. Berlin; 04/2018. Available from: URL: https://www.degam.de/files/Inhalte/Leitlinien-Inhalte/Dokumente/DEGAM-Praxisempfehlungen/PSA%20Screening/DEGAM%20PE_Hausaertztliche%20Beratung%20PSA-Screening_15-4-2018.pdf.
- Deutscher Bundestag, Beschlussempfehlung und Bericht des Ausschusses für Gesundheit. Brustkrebs – Mehr Qualität bei Früherkennung, Versorgung und Forschung – Für ein Mammographie-Screening nach Europäischen Leitlinien. Berlin; 2002 May 16. Bundestags-Drucksache 14/9122. <http://dip21.bundestag.de/dip21/btd/14/091/1409122.pdf>.
- Dierks ML, Schmacke N. Mammografie-Screening und informierte Entscheidung: mehr Fragen als Antworten. In: Böcken J, Braun B, Meierjürgen R, editors. *Gesundheitsmonitor 2014 – Bürgerorientierung im Gesundheitswesen*: Bertelsmann Stiftung – Barmer GEK. p. 55-91.
- Fachgremium Swiss Medical Board. Systematisches Mammographie-Screening. Zollikon: Swiss Medical Board; 2013. http://www.medical-board.ch/fileadmin/docs/public/mb/fachberichte/2013-12-15_bericht_mammographie_final_rev.pdf (30.07.2018).
- Felder S, Robra B-P. A preference-based measure for test performance with an application to prenatal diagnostics. *Stat Med* 2006; 25(21):3696-706.
- Felder S, Werblow A, Robra B-P. A priori risk and optimal test accuracy in prenatal diagnostics. *Med Decis Making* 2003; 23(5):406-13.
- Föry D. Kantone schaffen Früherkennungsprogramme für Brustkrebs ab; 2017. Available from: URL: <https://www.nzz.ch/schweiz/mammografie-kantone-schaffen-frueherkennungsprogramme-fuer-brustkrebs-ab-ld.1305648>.
- Frank L, Hueber S, Kühlein T, Schedlbauer A, Roos M. Zwischen Evidenz und Praxis: Wie bewerten Hausärzte Früherkennungsmaßnahmen? Eine Fragebogenstudie. *Z Evid Fortbild Qual Gesundhwes* 2018; 135-136:1-9.
- Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses: in der Fassung vom 18. Dezember 2008 veröffentlicht im Bundesanzeiger Nr. 84a (Beilage) vom 10. Juni 2009 in Kraft getreten am 1. April 2009, zuletzt geändert am 16. März 2018 veröffentlicht im Bundesanzeiger BANz AT 04.07.2018 B1 in Kraft getreten am 5. Juli 2018. Berlin: Gemeinsamer Bundesausschuss. Available from: URL: www.g-ba.de/downloads/62-492-1614/VerfO_2018-03-16_iK-2018-07-05.pdf.
- Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Richtlinie Ultraschallscreening auf Bauchaortenaneurysmen; 2016 Oct 20. Available from: URL: https://www.g-ba.de/downloads/40-268-4038/2016-10-20_US-BAA-RL_Erstfassung-RL_TrG.pdf.
- Gigerenzer G, Mata J, Frank R. Public knowledge of benefits of breast and prostate cancer screening in Europe. *J. Natl. Cancer Inst.* 2009; 101(17):1216-20. DOI: 10.1093/jnci/djp237
- Gummersbach E, in der Schmitt J, Mortsiefer A, Abholz H-H, Wegscheider K, Pentzek M. Willingness to participate in mammography screening – a randomized controlled questionnaire study of responses to two patient information leaflets with different factual content. *Dtsch Arztebl Int* 2015; 112(5):61-8.
- Habbema D, De Kok IM, Brown ML. Cervical cancer screening in the United States and the Netherlands: a tale of two countries. *Milbank Q* 2012; 90:5-37.
- Hakama M, Miller AB, Day NE, editors. *Screening for cancer of the uterine cervix*. Lyon: IARC; 1986. (IARC scientific publications; vol 76).

- Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. *N Engl J Med* 2016; 375(15):1415–24.
- Helou A. Krebsfrüherkennung im Nationalen Krebsplan. Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz 2014; 57:288–93.
- Hill GB, Spitzer WO, Ferenczi LZ. Optimal frequency of examinations aimed at detecting pre-symptomatic disease. *Can J Public Health* 1984; 75(6):419–24.
- Hubbard RA, Kulikowski K, Flowers CI, Yankaskas BC, Zhu W, Migliorette DL. Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. *Ann Intern Med* 2011; 155(8):481–92.
- Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Versicherteninformation zum Ultraschall-Screening auf Bauchaortenaneurysmen – Addendum zum Auftrag S13–04; 2016 Oct 13. IQWiG-Berichte 449. Available from: URL: www.iqwig.de/download/P16-01_Addendum_Versicherteninformation-zum-Ultraschall-Screening-auf-Bauchaortenaneurysmen.pdf.
- IQTIG. Zervixkarzinomscreening – Konzept einer gestuften Evaluation: Erstellt im Auftrag des Gemeinsamen Bundesausschusses. Berlin: Institut für Qualitätssicherung und Transparenz; 2017 Nov 30. Available from: URL: https://iqtig.org/downloads/berichte/2018/IQTIG_Evaluationskonzept-Zervixkarzinomscreening_Abschlussbericht_2018-04-09.pdf.
- Jatoi I, Baum M. Mammographically detected ductal carcinoma in situ: are we overdiagnosing breast cancer? *Surgery* 1995; 118(1):118–20.
- Jørgensen KJ. Mammography screening. Benefits, harms, and informed choice. *Dan Med J* 2013; 60:B4614.
- Jørgensen T, Jacobsen RK, Toft U, Aadahl M, Glümer C, Pisinger C. Effect of screening and lifestyle counselling on incidence of ischaemic heart disease in general population: Inter 99 randomised trial. *BMJ* 2014; 348:g3617.
- Kooperationsgemeinschaft Mammographie. Jahresbericht Evaluation 2015. Deutsches Mammographie-Screening-Programm. Berlin; Oktober 2017. Available from: URL: https://newsroom.mammo-programm.de/download/fachpublikation/KOOP-MAMMO_Jahresbericht_EVAL2015_20170926_web.pdf.
- Kötter T. DEGAM-Praxisempfehlung „Hausärztliche Beratung zu PSA-Screening“. *Z Allg Med* 2016; 92(12):495–499.
- LeFevre ML, on behalf of the U.S. Preventive Services Task Force. Screening for abdominal aortic aneurysm: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2014; 161(4):281–90.
- Lenz M, Buhse S, Kasper J, Kupfer R, Richter T, Mühlhauser I. Decision aids for patients. *Dtsch Arztebl Int* 2012; 109(22–23):401–8.
- Lühnen J, Albrecht M, Mühlhauser I, Steckelberg A. Leitlinie evidenzbasierte Gesundheitsinformation. Hamburg; 2017 [cited 2018 Sep 9]. Available from: URL: <http://www.leitliniegesundheitsinformation.de>.
- Makridakis S, Kirkham R, Wakefield A, Papadaki M, Kirkham J, Long L. Forecasting, uncertainty and risk; perspectives on clinical decision-making in preventive and curative medicine. *International Journal of Forecasting* 2018 (<https://www.sciencedirect.com/science/article/pii/S0169207017301346>).
- McNeil BJ, Weichselbaum R, Pauker SG. Fallacy of the five-year survival in lung cancer. *N Engl J Med* 1978; 299(25):1397–401.
- Meyer JMA, Swart E, Dierks M-L, Robra B-P. Wann werden empfohlene Mammabiopsien durchgeführt, und was erfährt der mammographierende Arzt darüber? *Der Pathologe* 1997; 18(1):60–6.
- Morabia A, Zhang FF. History of medical screening: from concepts to action. *Postgrad Med J* 2004; 80:463–469.
- Mühlhauser I. Zur Überschätzung des Nutzens von Prävention. *ZEFQ* 2014; 108:208–218.
- Murphy MF, Campbell MJ, Goldblatt PO. Twenty years' screening for cancer of the uterine cervix in Great Britain, 1964–84: further evidence for its ineffectiveness. *J Epidemiol Community Health* 1988; 42(1):49–53.
- Perrenoud B, Velonaki V-S, Bodenmann P, Ramelet A-S. The effectiveness of health literacy interventions on the informed consent process of health care users: a systematic review protocol. *JBI Database System Rev Implement Rep* 2015; 13(10):82–94.
- Ploug T, Holm S, Brodersen J. To nudge or not to nudge: cancer screening programmes and the limits of libertarian paternalism. *J Epidemiol Community Health*. 2012; 66:1193–6.
- Quanstrum KH, Hayward RA. Lessons from the Mammography Wars. *N Engl J Med* 2010; 363:1076–1079.
- Raftery J, Chorozioglou M. Possible net harms of breast cancer screening: updated modelling of Forrest report. *BMJ* 2011; 343:d7627.
- Reiff M, Ross K, Mulchandani S, Probert KJ, ScDy, Peyeritz RE, Spinner NB, Barbara A. Bernhardt BA. Physicians' Perspectives on the Uncertainties and Implications of Chromosomal Microarray Testing of Children and Families. *Clin Genet* 2013; 83:23–30.
- Robra B-P, Spura A. Versorgungsbedarf im Gesundheitswesen – ein Konstrukt. In: Klaber J, Geraedts M, Friedrich J, Wasem J, editors. Krankenhaus-Report 2018 – Schwerpunkt: Bedarf und Bedarfsgerechtigkeit. Stuttgart: Schattauer; 2018. p. 3–21.
- Robra BP, Swart E, Klemperer D. Überdiagnose und Übertherapie des Prostata-Karzinoms – ein unterschätztes Problem. In: Klaber J, Geraedts M, Friedrich J, Wasem J, editors. Krankenhaus-Report 2013: Mengendynamik: mehr Menge, mehr Nutzen? Stuttgart: Schattauer; 2013. p. 245–61.
- Scheff TJ. Decision rules, types of error, and their consequences in medical diagnosis. *Behavioral Science: Journal of the International Society for the Systems Sciences* 1963; 8(2):97–107.
- Schilling FH, Spix C, Berthold F, Erttmann R, Fehse N, Hero B et al. Neuroblastoma screening at one year of age. *N Engl J Med* 2002; 346(14):1047–53.
- Schmidt H. Eigenverantwortung und Darmkrebsvorsorge: warum und wie die Chronikerrichtlinie revidiert werden sollte. *Gesundheits- und Sozialpolitik* 2012; Heft 2: 45–55.
- Schmacke N, Richter P, Stamer M. Der weite Weg zur Partizipation. Hogrefe 2016.
- Seidel G, Münch I, Dreier M, Borutta B, Walter U, Dierks ML. Sind Informationsmaterialien zur Darmkrebsfrüherkennung in Deutschland verständlich oder verfehlen sie ihre Wirkung?:

- Bewertung von Gesundheitsinformationen durch Nutzerinnen und Nutzer mit unterschiedlichem Bildungshintergrund. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2014; 57(3):366–79.
- Sommer I, Titscher V, Gartlehner G. Participants' expectations and experiences with periodic health examinations in Austria – a qualitative study. BMC Health Serv Res 2018; 18(1):823.
- South-East London Screening Study Group. A controlled trial of multiphasic screening in middle-age: results of the South-East London Screening Study. Int J Epidemiol 1977; 6(4):357–63.
- Steckelberg A, Berger B, Köpke S, Heesen C, Mühlhauser I. Kriterien für evidenzbasierte Patienteninformationen. Z Arztl Fortbild Qualitatssich 2005; 99(6):343–51.
- Sprague BL, Stout NK, Schechter C, van Ravesteyn, Nicolien T, Cevik M, Alagoz O et al. Benefits, Harms, and Cost-Effectiveness of Supplemental Ultrasonography Screening for Women With Dense Breasts. Annals of Internal Medicine 2015; 162(3):157–66.
- Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, Holmes-Rovner M, Llewellyn-Thomas H, Lyddiatt A, Thomson R, Trevena L. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev. 2017 Apr 12; 4:CD001431.
- Stang A, Jöckel K-H. Einfluss des Krebs-Screenings auf die Gesamtmortalität: Welcher Effekt im günstigsten Fall erwartet werden kann. Dtsch Arztebl Int 2018; 115(29–30):481–6.
- Tabár L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grönroft O et al. Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. Lancet 1985; 1(8433): 829–32.
- Timmermans S, Buchbinder T. Patients-in-Waiting: Living between Sickness and Health in the Genomics Era. Journal of Health and Social Behavior 2010; 51:408–423.
- Tormann D. Das Geschäft mit der Angst – über die Pathologisierung der Lebensphasen von Frauen. clio 2013; (77):30–3. Available from: URL: https://www.gyn-bielefeld.de/sites/default/files/seiten_30-33_clio_77_final_web.pdf.
- Tosteson ANA, Yang Q, Nelson HD, Longton G, Soneji SS, Pepe M et al. Second opinion strategies in breast pathology: a decision analysis addressing over-treatment, under-treatment, and care costs. Breast Cancer Res Treat 2018; 167(1):195–203.
- Trevena L. Cancer screening – pros, cons, choice, and the patient. Aust Fam Physician 2009; 38:188–92.
- Walter U, Robra B-P, Schwartz FW. Prävention. In: Schwartz FW, Walter U, Siegrist J, Kolip P, Leidl R, Dierks M-L et al., editors. Public Health: Gesundheit und Gesundheitswesen. 3., völlig neu bearbeitete und erweiterte Auflage. München: Urban & Fischer in Elsevier; 2012. p. 196–223.
- Wegwarth O. Statistical Illiteracy in Residents: What They Do Not Learn Today Will Hurt Their Patients Tomorrow. J Grad Med Educ. 2013 Jun; 5: 340–341.
- Wegwarth O, Gigerenzer G. The Barrier to Informed Choice in Cancer Screening: Statistical Illiteracy in Physicians and Patients. Recent Results Cancer Res 2018; 210:207–21.
- Welch HG, Passow HJ. Quantifying the benefits and harms of screening mammography. JAMA Intern Med 2014; 174(3):448–54.
- Welch HG, Schwartz LM, Woloshin S. Overdiagnosed: Making people sick in the pursuit of health. Boston, Mass.: Beacon Press; 2011.
- Wellek S, Martus P, Wiegel T, Stöckle M. PREFERE aus der Retrospektive: Was waren die echten und was die angeblichen Probleme? Für ähnlich ambitionierte Studien lässt sich aus dem Scheitern eine Lehre ziehen. Deutsches Ärzteblatt 2017; 114(25):1243–4.
- Wilson JMG, Jungner G. Principles and practice of screening for disease. Public health paper 1968; No. 34. Geneva: World Health Organisation. Available from: <http://www.who.int/bulletin/volumes/86/4/07-050112BP.pdf>
- Wilt TJ, Jones KM, Barry MJ, Andriole GL, Culin D, Wheeler T et al. Follow-up of Prostatectomy versus Observation for Early Prostate Cancer. N Engl J Med 2017; 377(2):132–42.
- Windus G, Dierks M-L, Robra B-P, Swart E, Blank JH. Die Patientenaufklärung vor Früherkennungsuntersuchungen. Z Präventivmed Gesundheitsförd 1992; 4:16–19.
- Woods WG, Gao R-N, Shuster JJ, Robison LL, Bernstein M, Weitzman S et al. Screening of infants and mortality due to neuroblastoma. N Engl J Med 2002; 346(14):1041–6.
- Zahl P-H, Gøtzsche PC, Andersen JM, Mæhlen J. Results of the Two-County trial of mammography screening are not compatible with contemporaneous official Swedish breast cancer statistics. Dan Med Bull 2006; 53:438–40.



Prof. Dr. med. Bernt-Peter Robra, MPH

Epidemiologe und Sozialmediziner, von 1992 bis 2018 Direktor des Instituts für Sozialmedizin und Gesundheitsökonomie der Universität Magdeburg.



Prof. Dr. med. Norbert Schmacke

Arzt für innere Medizin, öffentliches Gesundheitswesen und Sozialmedizin. Hochschullehrer für Gesundheitswissenschaften am Institut für Public Health und Pflegeforschung der Bremer Universität. Von 2004 bis 2018 stellvertretendes unparteiisches Mitglied im Gemeinsamen Bundesausschuss.