

PERSPEKTIVEN

Screening – Früherkennung – Vorsorge

Standards der Bewertung von Screeningverfahren

Jürgen Windeler¹

¹ Fachbereich Evidenz-basierte Medizin, Medizinischer Dienst der Spitzenverbände der Krankenkassen (MDS), Essen.

Wie man den laufenden Diskussionen um die Etablierung eines Mammographiescreenings zur Früherkennung von Brustkrebs, der von Interessengruppen geforderten Neueinführung eines PSA-Screenings zur Früherkennung von Prostatakrebs oder den Diskussionen um die Früherkennung des Glaukoms, des Melanoms, von Morbus Alzheimer oder Diabetes mellitus entnehmen kann, erfreuen sich Screening bzw. Früherkennung ungebrochener Wertschätzung. Der Weisheit „Vorbeugen ist besser als Heilen“ folgend, erscheint vielen die Früherkennung im Sinne einer sog. Sekundärprävention als das mindestens zweitbeste Mittel dort, wo eine Primärprävention, das eigentliche Vorbeugen von Erkrankungen, auf Schwierigkeiten stößt.

Die grundsätzliche Wertschätzung solcher Maßnahmen trifft zusammen mit dem Umstand, dass es Früherkennung an einem negativen Feedback fehlt [7]: Alle Beteiligten sind vielmehr mit den Ergebnissen zufrieden. Ärzte sind froh, ihren Patienten etwas z.B. gegen die „Geißel“ Krebs anbieten zu können. Ist der Test positiv, hat man noch rechtzeitig etwas getan, ist der Test negativ, kann man sich erleichtert zurücklehnen. Aus Patientenperspektive ergibt sich eine ähnliche Sicht, die durch die Möglichkeiten einer frühen Therapie erweitert wird. Selbst auftretende Beeinträchtigungen und Nebenwirkungen dieser Therapie werden als im Sinne des wichtigen Ziels hinzunehmen toleriert. Sieht man von diesen unmittelbar spürbaren Beeinträchtigungen ab, so sind mögliche Schäden von Screeningmaßnahmen weder durch Patient noch durch Behandler unmittelbar erfahrbare. Auch der Nutzen nicht: Zwar mag ein Patient erleichtert sein, dass seine Erkrankung „rechtzeitig“ entdeckt

wurde, ob er oder sie aber durch die Teilnahme an einer Früherkennungsmaßnahme wirklich einen gesundheitlichen Vorteil hat, z.B. länger lebt, ist zumindest im Einzelfall nicht zu sagen. Da nur der positive Effekt plausibel erscheint, werden nachteilige Effekte nicht einmal in Erwägung gezogen und bleiben damit aus der Diskussion ausgeblendet.

Zusammengefasst ist die Diskussion um Screeningmaßnahmen von einer Vermutung positiver Effekte gekennzeichnet, die als so plausibel und hochrelevant eingeschätzt werden, dass eine sorgfältige, nüchterne Bewertung, eine sach- und datenbezogene Diskussion sowie eine angemessene Aufklärung der Patienten als nicht vorrangig und u.U. gar nicht zielführend angesehen werden, wenn man eine Steigerung der Teilnahmeraten z.B. an der Krebsfrüherkennung im Auge hat.

Tatsächlich ist jedoch die Nutzenvermutung keineswegs plausibel. Sicher ist im Gegenteil zunächst nur, dass Screeningmaßnahmen einen Schaden verursachen. Werden gesunde Personen einem diagnostischen Test unterzogen, so ist, ohne dass man etwas über den Nutzen der Anwendung dieses Tests etwa im Sinne einer Beeinflussung der Mortalität weiß, jedenfalls sicher, dass diese Patienten die Risiken der Testanwendung, die Risiken der Fehlklassifikationen sowie die Risiken potentieller Therapien tragen müssen, nicht zuletzt, dass sie eine vermutlich belastende Diagnose erfahren, die sie ohne Anwendung des Tests (noch) nicht belasten würde. Es ist in dieser Ausgangslage unmittelbar einsichtig, dass alle diese sicheren Belastungen zum einen quantifiziert werden müssten und zum anderen einem quantifizierten Nutzen gegenübergestellt werden. Nur der Beleg, dass eine positive Nutzen-Schaden-Relation besteht, rechtfertigt die Durchführung und die breite Einführung von Screeningmaßnahmen.

Ausgangspunkt der Bewertung

Die Ausgangssituation für Screeningmaßnahmen ist zunächst die Evidenz, dass bei wohl allen Erkrankungen frühe Stadien eine bessere „Prognose“ haben als späte Stadien, z.B. ausgedrückt in höheren 5-Jahres-Überlebensraten bei Frühstadien von Tumorerkrankungen gegenüber Spätstadien. Der zweite Aspekt dieser Ausgangssituation ist, dass ein diagnostisches Verfahren von ausreichender Einfachheit existiert, das die Frühstadien mit ausreichender Sicherheit (ausreichend hohe Sensitivität und Spezifität) zu entdecken in der Lage ist. Diese typische Konstellation führt zu der Schlussfolgerung, dass der Einsatz eines solchen diagnostischen Tests sinnvoll sei, um so die interessierenden Frühstadien zu entdecken mit der Folge, diese erfolgreicher therapieren zu können als die gleichen Patienten dann, wenn ihre Erkrankungen zu einem späteren Zeitpunkt diagnostiziert worden wären. Diese regelhafte Schlussfolgerung ist jedoch logisch falsch und empirisch mindestens voreilig. Der wesentliche Bruch in der Argumentation liegt darin, dass ein Vergleich der Prognosedaten verschiedener Patienten noch nichts über die *Änderung* der Prognose beim Übergang von einer Spätdiagnostik auf eine Frühdiagnostik *beim selben* Patienten aussagt. Um diese Frage der Prognoseänderung schlüssig zu beantworten, bedarf es geeigneter empirischer Untersuchungen. Wegen der absolut gesehen kleinen Effekte und der großen Latenz dieser Effekte müssen solche Studien sehr groß und sehr lang sein.

Neben dieser für die Einführung einer Früherkennungsmaßnahme vorletzten Frage (die letzte ist die nach der Machbarkeit, der Qualitätssicherung und der Akzeptanz der praktischen Durchführung) wird eine Reihe von expliziten Anforderungen an potentielle Screeningverfahren gestellt, die teilweise bereits in der bis-

herigen Argumentation enthalten sind. Aktuell sind diese z.B. auf der Internetseite des UK National Screening Committee nachzulesen (www.nsc.nhs.uk/pdfs/criteria.pdf, s. Kasten). Im Endeffekt gehen sie auf die in den 60er Jahren erarbeiteten WHO-Kriterien [11] zurück. Diese Anforderungen betreffen z.B. die Häufigkeit und Beeinflussbarkeit der in Frage stehenden Zielkrankheit, die Aussagekraft und Einfachheit des zu verwendenden Tests, die Existenz und die Effektivität zur Verfügung stehender Therapieverfahren sowie die Möglichkeit, einen diagnostischen Test als Früherkennungsmaßnahme in ein ganzes Programm einzubetten, wobei eben Fragen der Qualitätssicherung, der Anwendungsmodalitäten (Wiederholung), der Compliance, also generell der Organisation betrachtet werden müssen. Im Folgenden soll die Diskussion auf die Frage nach dem Nutzen, die gleichzeitig eine Reihe von gerade angesprochenen Aspekten umfasst, konzentriert werden. Die Illustration erfolgt unter Bezugnahme auf die Krebsfrüherkennung, neben der Untersuchung von Neugeborenen eines der bekanntesten Felder von Screeninguntersuchungen.

Ergebnisse zur Nutzenbewertung

Es ist für viele, insbesondere Krebskrankheiten belegt und unbestritten, dass die Anwendung von Screeningtests zu einer mehr oder weniger deutlichen Erhöhung der Zahl von Frühstadien führt. So wurde etwa in der tschechischen Studie zum Bronchialkarzinomscreening eine Rate von Frühstadien von 53% im Vergleich zu 21% in einer Kontrollgruppe ohne Screening gefunden [9]. Bei der Durchführung eines Neuroblastomscreenings stieg die Rate von Frühstadien ebenfalls um den Faktor 3 an [8]. Wie oben bereits ausgeführt, ist dies allein noch nicht als Nutzen von Tests anzusehen. Hinzu kommt, dass dieser Anstieg an Frühstadien überhaupt nur dann einen ersten Anhaltspunkt für einen potentiellen Nutzen bedeuten könnte, wenn er auf eine Verschiebung hindeutete, in dem Sinne, dass Erkrankungen, die „spontan“ in einem Spätstadium entdeckt werden,

The criteria for appraising the viability, effectiveness and appropriateness of a screening programme

All of the following criteria should be met before screening for a condition is initiated:

The condition

- 1.1. The condition should be an important health problem.
- 1.2. The epidemiology and natural history of the condition, including development from latent to declared disease, should be adequately understood and there should be a detectable risk factor, or disease marker and a latent period or early symptomatic stage.
- 1.3. All the cost-effective primary prevention interventions should have been implemented as far as practicable.

The test

- 1.4. There should be a simple, safe, precise and validated screening test.
- 1.5. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.
- 1.6. The test should be acceptable to the population.
- 1.7. There should be an agreed policy on the further diagnostic investigation of individuals with a positive test result and on the choices available to those individuals.

The treatment

- 1.8. There should be an effective treatment or intervention for patients identified through early detection, with evidence of early treatment leading to better outcomes than late treatment.
- 1.9. There should be agreed evidence-based policies covering which individuals should be offered treatment and the appropriate treatment to be offered.
- 1.10. Clinical management of the condition and patient outcomes should be optimised by all health care providers prior to participation in a screening programme.

The screening programme

- 1.11. There must be evidence from high quality Randomised Controlled Trials that the screening programme is effective in reducing mortality or morbidity. Where screening is aimed solely at providing information to allow the person being screened to make an "informed choice" (e.g. Down's syndrome, cystic fibrosis carrier screening), there must be evidence from high quality trials that the test accurately measures risk. The information that is provided about the test and its outcome must be of value and readily understood by the individual being screened.
- 1.12. There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/intervention) is clinically, socially and ethically acceptable to health professionals and the public.
- 1.13. The benefit from the screening programme should outweigh the physical and psychological harm (caused by the test, diagnostic procedures and treatment).
- 1.14. The opportunity cost of the screening programme (including testing, diagnosis, treatment, administration, training and quality assurance) should be economically balanced in relation to expenditure on medical care as a whole (i.e. value for money).
- 1.15. There must be a plan for managing and monitoring the screening programme and an agreed set of quality assurance standards.
- 1.16. Adequate staffing and facilities for testing, diagnosis, treatment and programme management should be made available prior to the commencement of the screening programme.
- 1.17. All other options for managing the condition should have been considered (e.g. improving treatment, providing other services), to ensure that no more cost-effective intervention could be introduced or current interventions increased within the resources available.
- 1.18. Evidence-based information, explaining the consequences of testing, investigation and treatment, should be made available to potential participants to assist them in making an informed choice.
- 1.19. Public pressure for widening the eligibility criteria for reducing the screening interval, and for increasing the sensitivity of the testing process, should be anticipated. Decisions about these parameters should be scientifically justifiable to the public.

durch Anwendung eines Screeningtests früher entdeckt werden. Man müsste also eine gleichzeitige Abnahme von Spätstadien beobachten. Im

Neuroblastombeispiel ist dies nicht der Fall. Die Rate der Spätstadien ist unverändert geblieben, der deutliche Anstieg der Frühstadien also nicht

PERSPEKTIVEN

Screening – Früherkennung – Vorsorge

Ausdruck einer Verschiebung, sondern von Überdiagnostik (s.u.).

Als deutlicheres Indiz für einen möglichen Benefit der Anwendung von Screeningmaßnahmen wird häufig die empirisch gestützte Beobachtung angesehen, dass Personen, deren Erkrankung durch Screening (früh) entdeckt wurde, eine längere mittlere Überlebenszeit haben, als Personen, deren Erkrankung durch Symptome (später) auffällig wurde. Diese Beobachtung ist jedoch aus zwei wohl bekannten Gründen nicht oder jedenfalls nicht eindeutig als Effekt des Screenings zu interpretieren.

Lead Time

Zum einen führt das sog. Lead-Time-Problem dazu, dass der offensichtliche Überlebensvorteil auch dann resultiert, wenn die in Frage stehende Erkrankung gar nicht beeinflussbar ist. Im hypothetischen Beispiel einer solchen Erkrankung, die typischerweise im Alter von 55 Jahren auftritt und an der die Patienten nach ca. 10 Jahren versterben, hätten die Patienten, die (spät) im Alter von 63 Jahren diagnostiziert würden, eine mittlere Überlebenszeit von 2 Jahren nach Diagnosestellung, diejenigen, die (früh) mit 57 Jahren diagnostiziert würden, eine mittlere Überlebenszeit von 8 Jahren nach Diagnosestellung, also ein sehr deutlicher Unterschied, der aber nicht als Vorteil für die zweite Patientengruppe interpretiert werden kann. Im Gegenteil würde man die 6 Jahre längere Kenntnis der (und Beeinträchtigung durch die) Diagnose eher als Schaden sehen. In dieser reinen Form ist das Interpretationshindernis durch die Lead Time (Lead-Time-Bias) natürlich nicht zu beobachten. Es ist jedoch davon auszugehen, dass jeder Vergleich von früh diagnostizierten und spät diagnostizierten Erkrankungen durch die Lead-Time-Problematik beeinträchtigt wird. Ein Vergleich der Prognose nach Diagnosestellung ist daher für die Nutzenbewertung von Screeningmaßnahmen untauglich. Nebenbei gilt das generell für den Versuch, den Behandlungs- oder Versorgungsfortschritt bei (Krebs-) Erkrankungen z.B. durch den Ver-

gleich von 5-Jahres-Überlebensraten zu messen [3].

Length-Bias

Das zweite zentrale Problem, das die Interpretation solcher Daten behindert, liegt im sog. Length-Bias, welches mit dem Problem der Überdiagnostik starke strukturelle Ähnlichkeiten hat. Es entsteht dadurch, dass durch wiederholte Screeninguntersuchungen eher langsam progredierende Erkrankungen entdeckt werden, die im Vergleich zu schnell fortschreitenden Verlaufsformen (die im Screeningintervall symptomatisch werden) eine bessere Prognose haben und dadurch, dass sie bevorzugt entdeckt werden, die Screeninggruppe insgesamt prognostisch günstiger erscheinen lassen als eine Gruppe ohne Screening. Auch hier ist also der Unterschied in der Prognose nicht auf eine Änderung der Prognose bei den Teilnehmern zurückzuführen.

Das Length-Bias resultiert aber nur aus einem spezifischen Selektionsmechanismus. Andere sind in dem gesamten Phänomen der Überdiagnostik zu sehen. Ein offensichtlicher Überlebensvorteil könnte darin begründet sein, dass die Anwendung des Screeningtests (aufgrund seiner speziellen Technik oder auch seiner Durchführungsmodalitäten) dazu führt, dass eine besondere Gruppe von Patienten diagnostiziert wird, die bisher nicht identifiziert worden ist. Zur Illustration stelle man sich z.B. vor, dass ein Screeningtest nur Patienten mit einer bestimmten Rezeptorkonstellation entdeckt, die bisher nicht entdeckt worden wären und deren Erkrankung einen eher gutartigen Verlauf nimmt, so dass sie nicht an ihrer Erkrankung versterben. Ein Vergleich von durch Screening entdeckten und durch Symptomatik auffällig gewordenen Patienten würde dann einen deutlichen Überlebensvorteil der ersten Gruppe zeigen, der aber nur darauf zurückzuführen ist, dass in den beiden Gruppen zwei unterschiedliche Erkrankungen mit unterschiedlicher Prognose betrachtet werden. Auch dieses Problem wird in der reinen Form nicht auftreten, man kann jedoch sicher davon ausgehen, dass unter den durch Screening entdeck-

ten Fällen solche sind, die ohne Screeningtest unerkannt geblieben wären und – wichtiger – durch dieses Nichterkennen auch nicht beeinträchtigt wurden.

Überdiagnostik führt zu einer verzerrten Bewertung von Screeningverfahren und zu gesundheitlichen Nachteilen für die Teilnehmer. Überdiagnostik ist daher eines der zentralen Probleme und ein zentrales Schadenspotential bei allen Screeningverfahren. Da die Überdiagnostik – dies ist das Selbstverständnis der Durchführung von Screeningmaßnahmen – immer eine Therapie und damit immer eine Übertherapie nach sich zieht, liegt die Belastung dieser Teilnehmer nicht nur in einer klinisch irrelevanten (nicht falschen!) Etikettierung, sondern eben auch in der Belastung durch Beeinträchtigungen und Risiken der angewandten Therapien. Für das Neuroblastom, das Prostatakarzinom und das Bronchialkarzinom, vermutlich auch für das Mammakarzinom, sowie speziell für diejenigen Tumorarten, für die das Screening Vorstadien identifizieren soll, also z.B. das Melanom oder das Dickdarmkarzinom, spielt die Problematik der Überdiagnostik für die Bewertung von Screeningmaßnahmen eine ganz zentrale Rolle. Geht man etwa davon aus, dass im fortgeschrittenen Alter die Mehrzahl aller Männer ein Prostatakarzinom in sich trägt, so wird die Problematik eines möglicherweise hoch sensitiven Screenings, das alle diese Tumoren entdecken könnte, sehr deutlich.

Es sei ergänzend darauf hingewiesen, dass natürlich alle Vergleiche von Screeningteilnehmern gegenüber Nichtteilnehmern die Probleme aller epidemiologischen Beobachtungsstudien aufweisen, die man unter dem Stichwort Selection-Bias zusammenfassen kann: Teilnehmer und Nichtteilnehmer werden sich in einer Reihe von Kriterien unterscheiden, von denen anzunehmen, jedenfalls nicht ausgeschlossen ist, dass sie mit der Krankheitsprognose korreliert sind. Sehr bedeutend ist in diesem Zusammenhang der mutmaßlich einflussreiche, aber schwer befriedigend zu operationalisierende Faktor des sozioökonomischen Status. Dieses Problem ist nicht spezifisch für die Bewertung

von Screeningmaßnahmen, sondern gilt für die Bewertung jedweder Intervention, siehe etwa die aktuelle Diskussion um die Hormonsubstitution in der Postmenopause. Unter anderem aus diesem Grund sind Versuche, Screeningmaßnahmen mit epidemiologischen Daten oder Studien (Mortalitätsverläufe für bestimmte Tumorarten, Fallkontrollstudien) zu bewerten, problematisch.

Vermeidung von Bias

Es gibt nach derzeitigem Stand der Erkenntnis nur ein aussagekräftiges Studiendesign, welches in der Lage ist, valide Aussagen zum Nutzen von Screeningmaßnahmen zu liefern, die randomisierte, prospektive Studie.

In einer solchen Studie wird die Einladung zur Screeningteilnahme mit der Nichteinladung oder, in einem anderen Design, nach Einverständnis zur Teilnahme die Teilnahme gegen die Nichtteilnahme randomisiert und verglichen. Der typische Ansatz ist dabei der erste, der zwar aufwendiger ist, aber valide Ergebnisse für die Praxis liefert. Im Bereich der Krebsfrüherkennung wurden in den 70er und 80er Jahren vier Studien zum Screening auf Bronchialkarzinom durchgeführt (Übersicht bei [9]), es liegen sieben solcher Studien für ein Mammographiescreening vor [6], vier Studien zum Haemocult®-Screening auf Dickdarmkarzinom [10] und eine abgeschlossene Studie zum PSA-Screening [4]. Die Studien umfassten 10 000 bis über 150 000 Teilnehmer, zwei noch laufende Screeningstudien schließen ebenfalls mehr als 100 000 Teilnehmer ein [1].

Solche Studien mögen von ihrem Aufwand her und der zeitlichen Verzögerung von Antworten unbefriedigend erscheinen. Keine Antworten zu haben sollte allerdings als noch unbefriedigender wahrgenommen werden.

Methodisch hochwertige Studien haben zwei entscheidende Pluspunkte: Durch eine randomisierte Vorgehensweise werden alle beschriebenen Bias-Probleme ausgeschaltet. Dies bedeutet natürlich nicht, dass es keine Lead Time und keine Überdiagnostik mehr gibt. Diese werden aber in den Vergleich einbezogen, so dass also die

Interpretation einer solchen Studie lauten würde: *Trotz* Lead Time und *trotz* Überdiagnostik kommt die Studie zu einem positiven Ergebnis. Neben dem völligen Beseitigen aller störenden Verzerrungsmöglichkeiten haben solche Studien aber einen weiteren großen Vorteil. Die Argumentation für eine Screeningmaßnahme setzt sich genau genommen aus einer längeren Argumentationskette zusammen. Für jedes Glied dieser Kette bräuchte man substantielle empirische Daten, um die Kette belastbar knüpfen zu können und so zu einer (durchaus möglichen) Entscheidung für eine Screeningmaßnahme ohne Ergebnisse aus randomisierten Studien zu kommen. Die randomisierte Studie dagegen betrachtet das Ergebnis der gesamten Kette, ohne sich um die einzelnen Glieder zu kümmern und kümmern zu müssen. Ein Beispiel mag dies verdeutlichen: Eines der zentralen Glieder in der Argumentationskette ist die Evidenz dafür, dass eine frühzeitige Therapie (die Therapie von Frühstadien) bei einem Patienten überhaupt effektiv ist, dass sie aber insbesondere effektiver ist als eine späte Therapie (Therapie von späteren Stadien) bei diesem gleichen Patienten. Dieser Beleg ist nicht dadurch erbracht, dass die Effektivität einer frühen Therapie gezeigt ist (die späte Therapie könnte genauso gut sein), und sie ist natürlich auch nicht dadurch erbracht, dass eine „übliche Therapie“ bei symptomatischen Patienten effektiv ist (diese könnte bei asymptomatischen Patienten weniger effektiv sein). Ausgehend von einem dieser beiden Belege würde die weitere Argumentation sehr weitgehende Daten zum Spontanverlauf der Erkrankung erfordern, zur Quantifizierung von Therapieeffekten bei verschiedenen Krankheitsstadien, und sie würde eine Reihe weiterer Annahmen (z.B. Nichtreversibilität der Erkrankung) erfordern, die man alle in einer randomisierten Studie nicht braucht. Das Gleiche gilt für das Ausmaß der Überdiagnostik und Übertherapie, für die Frage möglicher Schäden durch die Anwendung von Screeningverfahren und sogar für die Klassifikationsgüte des Tests (Sensitivität und Spezifität). Randomisierte Studien reduzieren also tatsächlich

den für eine Entscheidung notwendigen Daten- und Argumentationsaufwand.

Insgesamt sind für die Bewertung von Screeningmaßnahmen also randomisierte Studien von herausragender Bedeutung. Wie weit man mit dieser Forderung geht, ob man etwa wie in einer Publikation aus England [2] auch für das Neugeborenen-screening bevölkerungsweite Studien mit mehreren Millionen eingeschlossenen Kindern fordert, ist eine Frage der Abwägung, der Machbarkeit und der Akzeptanz. Die Neuroblastom-Studie ist ein Beispiel dafür, dass solche Studien, hier unter Einschluss von 4 Mio. Kindern, in Deutschland machbar und (selbstverständlich!) ethisch vertretbar sind. Da erfahrungsgemäß eine einigermaßen befriedigende Evaluation nach Einführung einer Screeningmaßnahme nicht mehr realisierbar ist, ist die Durchführung von aussagefähigen Evaluationsstudien vor Einführung notwendig.

Schlussbemerkung

Eine Frage ist jedoch, welche Bedeutung Ergebnisse solcher aufwendigen Evaluationen für den Diskussionsprozess über die Einführung von Früherkennungsmaßnahmen tatsächlich haben. Hier ergibt sich ein etwas zwiespältiges Bild.

Während ein Neuroblastomscreening seit Bekanntwerden der ungünstigen Ergebnisse nicht weiter verfolgt wird (diese Ergebnisse sind übrigens mit erstaunlicher Gelassenheit zur Kenntnis genommen worden, obwohl es um eine so sensible Thematik wie Krebs bei Kindern ging), zeigt die Debatte um die sieben randomisierten Studien zur Mammographie, dass die aus einer intensiven Analyse abzuleitende eher skeptische Haltung nicht erwünscht war. Noch krasser ist das Beispiel der einzigen randomisierten Studie zum PSA-Screening, die – mit einem klaren negativen Ergebnis – nirgends in der deutschsprachigen urologischen Literatur erwähnt oder jedenfalls angemessen gewürdigt wird. Dass die evidenzbasierte S3-Leitlinie bei Kenntnis, aber Nichterwähnung der Studie zu einer positiven Empfehlung dieser Früherkennungsmaßnahme kommt, zeigt, dass offen-

PERSPEKTIVEN

Screening – Früherkennung – Vorsorge

bar solche Art der Evidenz nicht immer geschätzt wird. Sie zeigt trivialerweise auch, dass man diese extrem aufwendigen Studien nicht machen sollte, wenn die Ergebnisse ohne praktische Relevanz bleiben. Wenn im Fall des Fehlens solcher Studien nicht klar und offensiv gegen eine Früherkennungsmaßnahme votiert wird, dann wäre von Befürwortern eine lückenlose Argumentationskette, deren einzelne Glieder sorgfältig mit Evidenz belegt sind, zu fordern, bevor positive Entscheidungen über Screeningverfahren gefällt werden. Denn, um noch einmal daran zu erinnern: „The unfortunate truth is that cancer screening will always result in harm, harm that would never have occurred in the absence of screening. Whether benefit outweighs the harm is of utmost importance when evaluating screening modality“ [5].

10. Towler B, Irwig L, Glasziou P, Kewenter J, Weller D, Silagy C. A systematic review of the effects of screening for colorectal cancer using the faecal occult blood test, Hemoccult. Br Med J 1998;317: 559–65.
11. Wilson JMG, Jungner G. Principles and practice of screening for disease. Genf: WHO, 1968.

Literatur

1. De Koning HJ, Auvinen A, Sanchez AB, da Silva FC, Ciatto S, Denis L, Gohagen J, Hakama M, Hugosson J, Krause R, Nelen V, Prorok PC, Schröder FH. Large-scale randomized prostate cancer screening trials: program performances in the European randomized screening for prostate cancer trial and the prostate, lung, colorectal and ovary cancer trial. Int J Cancer 2002;97:237–44.
2. Dezateux C. Evaluating newborn screening programmes based on dried blood spots: future challenges. Br Med Bull 1998;54:877–90.
3. Donnelly B. Do increased 5-year survival rates in prostate cancer indicate better outcomes? JAMA 2000;284:2053–5.
4. Labrie F, Candas B, Dupont A, Cusan L, Gomez JL, Suburu RE, Diamond P, Levesque J, Belanger A. Screening decreases prostate cancer death: first analysis of the 1988 Quebec prospective randomised controlled trial. Prostate 1999;38:83–91.
5. Marcus PM. Lung cancer screening: an update. J Clin Oncol 2001;19:Suppl:83s–6s.
6. Olsen O, Gotzsche PC. Cochrane review on screening for breast cancer with mammography. Lancet 2001;358:1340–2.
7. Ransohoff DF, McNaughton Collins M, Fowler FJ. Why is prostate cancer screening so common when the evidence is so uncertain? A system without negative feedback. Am J Med 2002;113: 663–7.
8. Schilling FH, Spix C, Berthold F, Erttmann R, Fehse N, Hero B, Klein G, Sander J, Schwarz K, Treuner J, Zorn U, Michaelis J. Neuroblastoma screening at one year of age. N Engl J Med 2002;346:1047–53.
9. Strauss GM. Randomized population trials and screening for lung cancer. Cancer 2000;89: Suppl:2399–421.

*Korrespondenzanschrift**Prof. Dr. Jürgen Windeler**Fachbereich Evidenz-basierte Medizin
Medizinischer Dienst der Spitzenverbände
der Krankenkassen (MDS)**Lützowstraße 53**45141 Essen**Telefon (+49/201) 8327-125**Fax -402**E-Mail: j.windeler@mds-ev.de*