

# Mammografie-Screening: Falsche Befunde und das Spiel mit geschätzten Zahlen

Ulrich Placzek

*Das vom Deutschen Bundestag 2002 parteiübergreifend als Investition in die Zukunft beschlossene Mammografie-Screening-Programm steht zunehmend wegen des ungünstigen Nutzen-Risiko-Verhältnisses in der Kritik: Die falsch positiven und die falsch negativen Befunde sind immer mehr in den Fokus des Interesses geraten. Es ist nicht hinnehmbar, dass bei einem derart aufwändigen Programm etwa jedes vierte Karzinom nicht gefunden wird; nicht hinnehmbar ist ebenso, dass bei einer großen Zahl von Überdiagnosen Frauen auch zukünftig unnötig operiert, bestrahlt und chemo- bzw. hormontherapiert werden.*

Die Zweifel der Screening-Skeptiker werden dadurch befeuert, dass die systemimmanenten Schwächen des Screenings mit geschätzten Zahlen und mit Halbwahrheiten kleingeredet werden und dass sich das deutsche Mammografie-Screening wegen der konstant niedrigen Teilnehmerrate mit einem sogenannten Paradigmenwechsel von den Europäischen Leitlinien verabschiedet hat: „Informierte Inanspruchnahme“ wird jetzt als entscheidend angesehen.

## Frühe Entdeckung

Maligne Zellen können jahrelang ohne jede Aktivität in Geweben liegen (tumor dormancy). Bei eintretendem Wachstum kommt es ab einer kritischen Tumorgöße zu Sauerstoffmangel, der eine angiogene Kaskade auslöst: Endothelzellen migrieren, ein Netzwerk von neuen Gefäßen entsteht, die neuen Blutgefäße wachsen in den Tumor ein und stellen damit eine optimale Voraussetzung für schnelles Tumorstadium und Metastasierung dar – die invasive Karriere des Tumors kann beginnen.

Dieser als „angiogenic switch“ bezeichnete Effekt (14) erfolgt bei einer Tumormasse von etwa  $10^6$  Tumorzellen, was einer Tumorgöße von

1–2 mm entspricht, also zu einem Zeitpunkt, vor dem die Diagnose lediglich mittels Stammzellen gestellt werden könnte. Bildgebende Verfahren kommen in dieser Hinsicht als Früherkennungsmaßnahme regelmäßig zu spät. Stammzellen sind dagegen bereits vor dem Angiogenetic switch auffällig (c-Met, Oct 3–4, ZEB1, CD44 und c-Myc), nämlich dann, wenn eine Tumorzelle pro eine Million Monozyten vorkommt (30) (s. Tab.).

Bezogen auf bildgebende Verfahren ist die Zeit für ein radikales Umdenken gekommen, da sich die fundamentale Philosophie, die eine frühe Entdeckung des Tumors ins Zentrum gerückt hat, definitiv als Fehler herausgestellt hat (19). Und: Die durch Bildgebung gefundenen kleinen Karzinome sind nicht unbedingt früh erkannt, es sind eher die harmlosen und langsam wachsenden Tumoren.

Der Deutsche Bundestag formuliert im Nationalen Krebsplan 2008 das Ziel, Brustkrebskrankungen in einem prognostisch günstigen Stadium zu entdecken, um damit den Erfolg einer Therapie und die Überlebenschancen der betroffenen Frauen zu erhöhen; hierdurch könne die Brustkrebssterblichkeit gesenkt werden. Mit dieser selbst 2014 noch vorgetragenen Aussage (5) unterliegt die Bundesregierung einer dem epidemiologischen Basiswissen zuzuordnenden statistischen Verzerrung (Length Time Bias), die inzwischen auch von den politisch Verantwortlichen als solche verstanden werden sollte (27).

## Teilnehmerrate und Brustkrebsmortalität

In Deutschland wurde das Screening deshalb eingeführt, weil in acht randomisierten internationalen Studien die Reduktion der Brustkrebsmortalität mit 20–30% angegeben wurde – hierbei wurde die weltweit beachtete dänische Studie unabhängiger Autoren, die zu ganz anderen Ergebnissen gekommen ist, allerdings nicht berücksichtigt (5, 12).

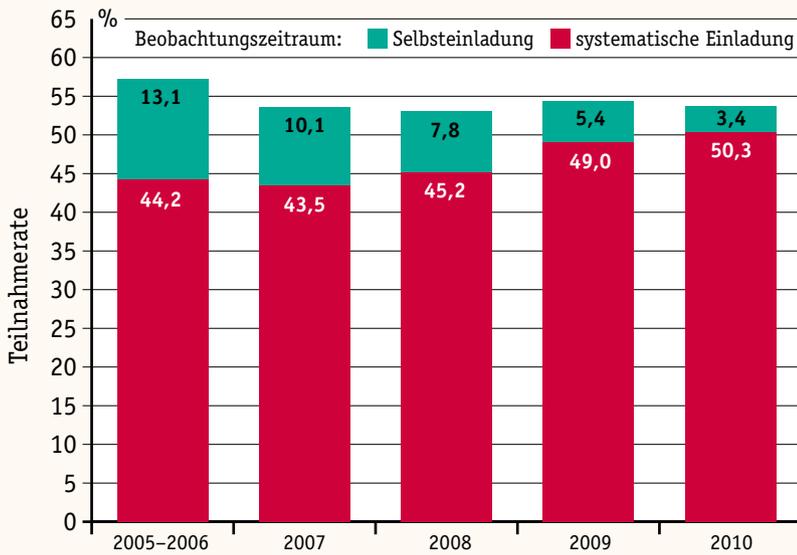
Gebendet durch von Screening-Protagonisten genannte „relative“ Zah-

### Zeitpunkt der Brustkrebsfrüherkennung anhand von Zellzahl und Tumorgöße

| Tumor       | Stammzellen | NMR, CT, PET | Mammografie, Ultraschall | Selbstuntersuchung |
|-------------|-------------|--------------|--------------------------|--------------------|
| Durchmesser | 1–2 mm      | 5–7 mm       | 6–8 mm                   | 20 mm              |
| Zellzahl    | $10^6$      | $10^7$       | $10^8$                   | $10^9$             |

Tab.: Bildgebende Verfahren kommen als Früherkennungsmaßnahme regelmäßig zu spät, Stammzellen sind dagegen bereits vor dem angiogenetic switch auffällig.

## Teilnehmerraten am Brustkrebs-Screening in Deutschland



nach (23)

Abb. 1: Die konstant niedrige Teilnehmerrate verfehlt die von der Europäischen Leitlinie vorgegebenen 70–75 % dramatisch.

len verhinderteter Todesfälle wurde damals nicht gesehen, dass die durch Screening zu erwartende Reduktion der Brustkrebssterblichkeit korrekt mit etwa 0,1 % hätte angegeben werden müssen (absolute Reduktion). Die nach neun Jahren Screening konstant niedrige Teilnehmerrate war bereits 2006 vorhersehbar (26). Sie verfehlt die von der Europäischen Leitlinie (7) vorgegebenen 70–75 % dramatisch (s. Abb. 1) und es ist beim besten Willen nicht zu erkennen, wie sich das in absehbarer Zeit ändern könnte. In Deutschland stimmt die Hälfte der eingeladenen Frauen zwischen dem 50. und 69. Lebensjahr mit den Füßen gegen das Screening; ein noch so gutes Screening-Programm könnte schon aus diesem Grund nur sehr begrenzt Einfluss auf die Brustkrebssterblichkeit nehmen.

Befürworter und Gegner des Screenings argumentieren seit Jahrzehnten mit den immer gleichen Studien, von denen die älteste vor 50 Jahren in New York und die jüngste vor 23 Jahren in Großbritannien begonnen wurde, ohne dass diese Studien die heutigen Therapieerfolge würdi-

gen können: Langzeiterfahrungen mit Tamoxifen oder Trastuzumab gab es zu diesem Zeitpunkt noch nicht.

Die Studien von damals bilden den Zeitraum von der „Antike bis zum Mittelalter“ der Brustkrebsdiagnostik und -behandlung ab (2). In der „Neuzeit“ angekommen, wurden in den letzten Jahren große Studien vorgelegt, die hinsichtlich der Brustkrebsmortalität einen Nulleffekt (20, 31) zeigen. Das Swiss Medical Board hat aus genau diesem Grund empfohlen, das Mammografie-Screening in der Schweiz nicht einzuführen (2). Die Arbeitsgruppen aus Trondheim (32) und Rotterdam (9) haben von einem deutlichen Screening-Erfolg berichtet: Vor Einführung des Screenings in den Niederlanden (1986–88) nahm die Mortalität jährlich um 0,3 % zu, mit Einführung des Screenings fiel sie jährlich um 1,7 % ab. Die Reduktion der Mortalität in 20 Jahren Screening gaben die Autoren mit relativen 31 % an, wobei dieser Erfolg ebenso dem Screening (15,7 %) wie auch einer verbesserten adjuvanten Therapie (13,9 %) zugeschrieben wurde. Mit Spannung wartet man bei einer solch

konkreten Aussage mit wissenschaftlichem Anspruch darauf, wie diese Zahlen ermittelt werden konnten: Sie wurden in Rotterdam mittels eines Mikrosimulation-Computer-Modells und von der norwegischen Gruppe mittels einer Computersimulation geschätzt.

Kritisch ist anzumerken, dass Screening-Effekte niemals sofort, wie von der Rotterdam-Gruppe angegeben, sondern erst nach einem Follow up von 10 Jahren (5) beurteilt werden dürfen und dass Schätzungen nicht als wissenschaftlich fundiert, sondern als von individueller Interpretation geprägt anzusehen sind. Nach Kenntnis der Bundesregierung erlaubt die wissenschaftliche Datenlage derzeit keine sichere Unterscheidung zwischen den Effekten des Screenings und den Effekten, die durch verbesserte Behandlungsmöglichkeiten erreicht werden können. Vom Bundesamt für Strahlenschutz beauftragt, genau diesen Punkt zu untersuchen, vermutet Hense, dass die Effekte auch in Zukunft nicht voneinander getrennt werden können (22).

Somit stellt sich ethisch und auch haftungsrechtlich die Frage, ob es unter den gegebenen Umständen verantwortbar ist, Frauen weiterhin mit kleingeredeten Screening-Risiken wie Überdiagnosen, die zu Übertherapie führen, zu belasten.

### Haftungsrecht: Kennzahlen und informierte Inanspruchnahme

Die Kooperationsgemeinschaft Mammografie wird von den gesetzlichen Krankenkassen und der Kassenärztlichen Bundesvereinigung getragen und ist mit der Koordination, Qualitätssicherung und Evaluation des Screenings betraut. Um die zwischen Screening-Befürwortern und -Kritikern kontrovers geführte Diskussion über die Größe des Nutzens und der Risiken in einen Konsens zu überführen, hat sie 2010 einen Wissenschaftsjournalisten beauftragt, „Kennzahlen“ zu

erstellen, die konkrete Aussagen zu den wesentlichen Fragen des Programms erlauben (16, 33).

Die vorgelegten Kennzahlen wurden nicht mathematisch ermittelt, sondern abgeschätzt. Dabei sollte gelten: Jede Kennzahl für sich soll plausibel sein, wichtige Verhältnisse sollen gewahrt bleiben (vermeidene Brustkrebs-Todesfälle zu Todesfällen wegen Brustkrebs) und jede Kennzahl soll nach Möglichkeit durch 5 teilbar sein. Diese Kennzahlen stellen ein Good-Case-Szenario dar, „sie vermitteln ein tendenziell optimistisches, aber noch realistisches Bild“ – dieser Satz für sich genommen ist schon eine intellektuelle Herausforderung. Eine hier exemplarisch herausgegriffene Kennzahl (für Überdiagnosen und Übertherapien) wurde, bezogen auf 1.000 Frauen und 20 Screening-Jahre, mit 5 ermittelt, wobei in der Hälfte der zitierten Studien zu dieser Frage gar keine Aussage gemacht wurde (die beiden Studien der Kooperationsgemeinschaft gehören dazu). Zwei Autoren (Götzsche und Mühlhauser) gaben die Zahl 10 an, ein Autor 2–10 und je ein weiterer Autor 5 bzw. 4; wie aus diesen Zahlen die Kennzahl 5 ermittelt wurde, erschließt sich rechnerisch nicht.

Nach heftiger Kritik an der Informationsbroschüre versuchen jetzt der Gemeinsame Bundesausschuss und die Kooperationsgemeinschaft mit emsiger Betriebsamkeit eine vermeintlich hochkarätige Diskussion darüber zu initiieren, „wie eine Information inhaltlich und gestalterisch aufbereitet werden muss, damit sie von möglichst vielen verstanden wird (...) In mehreren wissenschaftlichen Projekten und Gremien wird dieser Aufgabe nachgegangen, (...) wir arbeiten einerseits auf der wissenschaftlichen Ebene mit Experten zusammen, andererseits setzen wir vorliegende Erkenntnisse bereits um“ (17) – das mutet wie Lockendreher auf einer Glatze an. Ganz einfach und deutlich kostengünstiger wäre es, die ausgewogene dänische

Informationsbroschüre eins zu eins zu übernehmen (35).

Die Bundesregierung lässt wissen, dass sie nicht wertend zu einzelnen medizinisch-wissenschaftlichen Publikationen, deren Ergebnissen und Qualität Stellung nimmt; sie teilt aber dennoch mit, dass sie die kritischen Aussagen des Deutschen Netzwerks Evidenzbasierte Medizin nicht nachvollziehen kann und dass der Bericht des Swiss Medical Board nur eine von vielen aktuellen Veröffentlichungen ist.

Haftungsrechtlich ist zu sagen, dass eine Rechtsgrundlage für das Mammografie-Screening nur dann gegeben ist, wenn nach dem derzeitigen medizinischen Wissensstand der Nutzen die Risiken überwiegt. Falls die absolute Risikoverminderung 0,1 % beträgt, nehmen 999 gesunde Frauen solidarisch einen scheinbar kleinen Nachteil wie Überdiagnosen in Kauf, damit einer symptomlos Erkrankten das Leben gerettet werden kann. Bei dieser Sichtweise wird aber übersehen, dass Risiko unteilbar ist (21) und dass der Nachteil bei Überdiagnose überhaupt nicht klein ist. Deshalb muss die Patientin in einem persönlichen Gespräch mit dem Arzt von dem zu vermutenden Irrtum befreit werden, dass die Mammografie ausschließlich Nutzen hat. Sie muss wissen, wie wahrscheinlich falsch negative und falsch positive Befunde sind. Nur die informierte Einwilligung (informed consent) qualifiziert Frauen für eine selbstbestimmte Entscheidung (21). Auch eine eventuell überarbeitete Informationsbroschüre kann das Gespräch nur unterstützen und begleiten, ersetzen kann sie das Gespräch nicht.

### **Überdiagnosen: Prävalenz, positiv prädiktiver Wert, Bayes-Theorem**

Um diagnostische Verfahren sinnvoll einsetzen zu können, müssen zunächst Sensitivität und Spezifität bekannt sein. Bei der Interpretation der Ergebnisse muss die Prävalenz der vermu-

teten Erkrankung berücksichtigt werden, um den prädiktiven Wert möglichst realistisch abschätzen zu können. Kurvendiagramme erleichtern die formale Interpretation von Testergebnissen nach dem Bayes-Theorem (25).

Während die Sensitivität die Wahrscheinlichkeit für ein positives Testergebnis unter der Bedingung, tatsächlich erkrankt zu sein, beschreibt, drückt die Spezifität die Wahrscheinlichkeit für ein negatives Ergebnis unter der Bedingung aus, nicht erkrankt zu sein. Der Cut-off-Wert definiert das Verhältnis der beiden zueinander: Verschiebt man ihn zugunsten der Spezifität, verschlechtert man die Sensitivität und umgekehrt. Konkret auf das Mammografie-Screening bezogen heißt das, dass die ambitioniert möglichst niedrig gehaltene Zahl der nach der Mammografie wiedereinbestellten Frauen (Spezifität) zwangsläufig zu einer Erhöhung der Intervallkarzinome (Sensitivität) führen muss – es besteht also an dieser Stelle gar kein Grund, die hohe Qualität des Screenings zu preisen.

Verbesserungspotenziale (Gespräch, klinische Untersuchung und zusätzliche Mammasonografie) können von den systemimmanenten Schwächen der Mammografie (Brustdichte (24)) abgegrenzt werden. Screening-Ärzte dürfen sich deshalb nicht nur das gut honorierte Filetstück Mammografie aus der Diagnosekette der Brustkrebsfrüherkennung herauschneiden, sie müssen auch – trotz des damit verbundenen Zeitaufwands – bereit sein, mit Patientinnen zu reden und sie klinisch zu untersuchen. Das Arzt-Patienten-Gespräch und die klinische Untersuchung sind auch in Zeiten apparativ dominierter Medizin unverzichtbar. Sie zählen zum ärztlichen Selbstverständnis, sie tragen wesentlich zur angestrebten Qualität und Sicherheit von Früherkennung bei und sie könnten auch die Compliance für das Screening verbessern.

Der prädiktive Wert eines Testverfahrens erlaubt, den zunächst unbekann-

ten Status krank oder nicht krank zu antizipieren. Er hängt nicht nur von der Sensitivität und Spezifität, sondern ganz wesentlich von der Prävalenz ab: Die Prävalenz beschreibt die Wahrscheinlichkeit für das Vorliegen eines Merkmals und erlaubt, die Sinnhaftigkeit von Testverfahren vorweg zu beurteilen. Die Prävalenzabhängigkeit des positiven und des negativen prädiktiven Werts (ppV bzw. npV) wird durch das Bayes-Theorem beschrieben und lässt sich grafisch abbilden (s. Abb. 2).

Da Screening-Verfahren grundsätzlich im Niedrigprävalenzbereich liegen (bei der Screening-Mammografie liegt sie unter 1%), werden zwar etwas mehr Karzinome erkannt, sie werden aber wegen des schlechten ppV mit einem deutlichen Anstieg falsch positiver

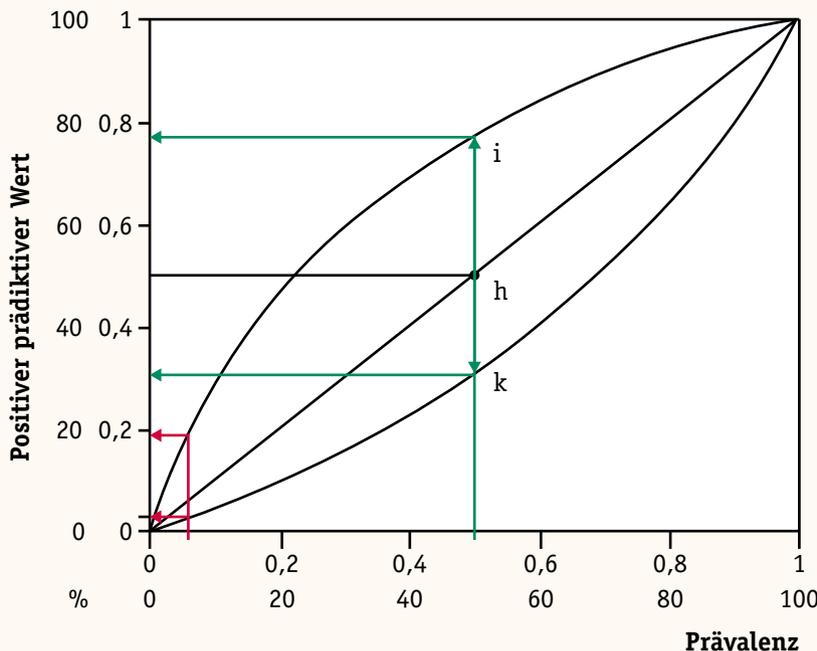
Befunde erkaufte. Die Frage nach der Verhältnismäßigkeit muss auch unter monetären Aspekten betrachtet werden. Ganz entscheidend muss aber das Schicksal der Frauen in den Blick genommen werden, die aufgrund der Krebsdiagnose operiert und multimodal adjuvant behandelt werden, obwohl viele von ihnen von ihrer Krebserkrankung nie in ihrem Leben erfahren hätten. Wir dürfen heute davon ausgehen, dass in wenigen Jahren das individuelle genetische Risikoprofil bestimmt werden kann (29). Frauen mit einem hohen genetischen Mammakarzinom-Risiko würden in einem eher mittleren Prävalenzbereich liegen. Diese Frauen haben wirklich einen Vorteil von der Mammografie.

Überdiagnosen können per se nicht als falsch positive Befunde bezeich-

net werden, jedenfalls nicht nach radiologischen oder histopathologischen Kriterien; falsch und für viele der betroffenen Frauen tragisch sind allerdings die aus vielen Diagnosen gezogenen Schlüsse, die zu Übertherapien führen. In den Ohren betroffener Frauen müssen Bemerkungen wie „Überdiagnosen sind tolerierbar und halten sich in Grenzen“ (9), sie „sprechen nicht gegen das Screening“ (3) zynisch klingen. Was es heißt, eine Krebspatientin zu sein, berichtet eine Betroffene so (13): Sie erwartete, dass sie als Screening-Erfolg eingestuft würde; ihre Erfahrung sagte das Gegenteil, das Screening hatte ihr beträchtliche Nachteile eingebracht: zwei große Exzisionen, eine Teilmastektomie, eine Rekonstruktion, mehrere Vollnarkosen, fünf Wochen Bestrahlung, chronische Infektionen, mehr als ein Jahr Arbeitslosigkeit und dann noch die Belastung ihrer Familie, besonders ihrer Tochter. Bei aller Empathie des Lesers dieser Zeilen muss gesagt werden, dass weder die Patientin noch die behandelnden Ärzte – und das weder ex ante noch ex post – wissen können, ob die Betroffene wirklich die von ihr gefühlten Nachteile oder gar lebensrettende Vorteile durch das Screening erfahren hat: Genau diese Tatsache charakterisiert das Problem mit den Überdiagnosen. Nach heutigem Kenntnisstand bleibt bei Diagnosestellung den Ärzten nur, die adäquate Therapie zu empfehlen. Nur das „Recht auf Nichtwissen“, also die Nichtteilnahme am Screening, kann vor den von der Patientin vermuteten Nachteilen bewahren.

Über Beschwerden nach Brustkrebsbehandlung gaben betroffene Frauen in einem Fragebogen Auskunft: 91% waren mit dem OP-Ergebnis zufrieden oder sehr zufrieden; 34% hatten Schmerzen im Brustbereich; 35% berichteten über eine eingeschränkte Schulter-/Armfunktion; jüngere Patientinnen litten stärker als ältere unter psychischen Folgen; 25% erlebten eine Veränderung in der Partnerschaft; 9% kontaktierten vorher, 19%

**Bayes-Theorem**



nach (25)

Abb. 2: Die Senkrechte auf der x-Achse schneidet bei einer mittleren Prävalenz von 50% in i die Kurve der positiven Testergebnisse, der zugehörige ppV wird mit 76% abgelesen (grüne Linien); die Wahrscheinlichkeit für den Status krank ist nach einem positiven Test um 26 Prozentpunkte angestiegen. Ergibt sich bei gleicher Prävalenz ein negativer Test, lässt sich über den Schnittpunkt k der npV 32% ablesen, die Wahrscheinlichkeit für den Status krank ist von 50 auf 32, also um 18 Prozentpunkte gefallen. Bei einer angenommenen Prävalenz von 5% (rote Linien) dagegen wäre der Zugewinn an Informationen, wie dem Diagramm zu entnehmen ist, deutlich geringer (12 Prozentpunkte). Der Informationszugewinn ist bei sehr niedrigen und sehr hohen Prävalenzwerten gering, am höchsten ist er bei mittleren Prävalenzwerten, wo sich die Kurven am weitesten von der Winkelhalbierenden entfernen (25).

nachher einen Psychiater/Psychotherapeuten; 14 % nahmen vorher, 26 % nachher Psychopharmaka ein (8).

Überdiagnosen sind nicht zählbar, sie können nur in Modellrechnungen, Extrapolationen oder mittels der Brustkrebssterblichkeit über Jahre indirekt abgeschätzt werden. Die individuell angegebenen Werte streuen deshalb erwartungsgemäß weit, zwischen 1–30 %. Vermutlich ist für Frauen weniger die absolute Zahl der Überdiagnosen relevant als vielmehr das Verhältnis von verhinderten Brustkrebstodesfällen zu Überdiagnosen: bei Gøtzsche liest man 1:10, bei anderen Autoren 1:1 (33).

### Intervallkarzinome

Intervallkarzinome sind Karzinome, die im Intervall von zwei Screening-Untersuchungen auftreten; sie werden, falls sie klinisch manifest werden, oftmals von Frauen selbst entdeckt. Die Gruppe aus Nijmegen gab 1999 Intervallkarzinome mit 30 % an (6); 2012 wurden in Nordrhein-Westfalen 22–23 % gefunden, 40 % der T2–T4-Tumore waren allerdings Intervallkarzinome (15). Die weitere Analyse der niederländischen Autoren ergab in 50 % „echte“ Intervallkarzinome, in 30 % Karzinome mit minimalen Malignitätszeichen, Untersuchungsfehler wurden mit 10–20 % angegeben, in 5 % lagen okkulte Karzinome bzw. technische Fehler vor.

Für die Qualitätsbeurteilung des Screenings sahen die Autoren die Analyse der Intervallkarzinome als unverzichtbar an (6). Aus ähnlichen Gründen ist in den Europäischen Leitlinien von 2005 zu lesen: „The purpose of monitoring interval cancers is twofold ... it is crucial since it serves both quality assurance and training ... it allows the calculation of parameters providing an early estimate of the impact of screening program ...“ (7).

Unklar bleibt, warum das Screening in Deutschland dieser Vorgabe der

Europäischen Leitlinie nicht folgt, wo sich Politik und Kooperationsgemeinschaft doch sonst in allem auf die europäischen Leitlinien berufen. Die seit Jahren wiederholte Aussage, dass es in Deutschland keine Krebsregister gebe, überzeugt nicht, da sogar von der Kooperationsgemeinschaft im Evaluationsbericht 2010 seit Jahren funktionierende Krebsregister in allen Bundesländern außer in Baden-Württemberg und Hessen angegeben werden (23). Als Zumutung ist die Aussage zu bewerten: „Bei wie vielen Frauen möglicherweise ein ‚Intervallkarzinom‘ vorliegt, ist bislang für Deutschland aus den alleinigen Angaben für Nordrhein-Westfalen nicht ablesbar. Hier warten wir auf das gesamtdeutsche Ergebnis, das erst nach vollständiger Aufstellung der Landeskrebsregister möglich sein wird“ (17). Bei dem enormen Aufwand, der um das Mammografie-Screening in Deutschland betrieben wird, wäre es schon längst möglich gewesen, entsprechende Daten in allen Bundesländern, außer den beiden genannten, zu erheben. 70–75 % der in Deutschland erfassten Fälle sollten auch hier genügen.

Unverständlich ist, dass ein Epidemiologe aus der Tatsache, dass die Zahl der Intervallkarzinome in Nordrhein-Westfalen mit 22 % angegeben wird, folgert, dass die Sensitivität der Screening-Mammografie 78 % betrage (18). Er unterliegt damit einem klassischen Trugschluss, da ja nicht alle Karzinome, die in einer Screening-Runde nicht gefunden werden, in den folgenden zwei Jahren als Intervallkarzinome auftauchen. Schon die bereits erwähnten nicht gefundenen Tumore mit minimalen Karzinomzeichen müssen zwar der Sensitivität angelastet werden; der wahrscheinlich größere Teil wird wohl erst in der folgenden Screening-Runde erkannt werden. Auch nehmen nicht alle Frauen nach zwei Jahren wieder am Screening teil; sie könnten mit einem nicht gefundenen Karzinom bei „tumor dormancy“ unbehelligt und glücklich weiterleben, ohne dass das

vom Screening-Programm erfasst würde. Die Sensitivität liegt folglich unter den angegebenen 78 %.

### Interessengemeinschaft Mammografie-Screening

Es fällt auf, dass von führenden Screening-Befürwortern ganz offensichtlich falsche Angaben gemacht werden und mit Halbwahrheiten argumentiert wird, und das in einer Fülle, dass es schwerfällt, jedem einzelnen Punkt zu widersprechen. Es hat sich eine Interessengemeinschaft um das Screening-Programm geschart, wobei den Screening-Ärzten die Kooperationsgemeinschaft als PR-Instrument dient. Screeningnahe Epidemiologen, die, wenn es um die Vorzüge des Screenings geht, wider besseres Wissen immer noch mit den großen relativen Zahlen argumentieren (10) und kleine Karzinome gebetsmühlenartig als früh erkannt betrachten, präsentieren sich auch durch verwirrende Rechenspiele und individuelle Schätzungen als Vorzeigewissenschaftler. Es bedarf eines erheblichen investigativen Aufwands, sich in diesem Wirrwarr halbwegs zurechtzufinden.

Interessenkonflikte bestehen bereits, wenn das Risiko der Beeinflussung besteht, und nicht erst, wenn eine solche Beeinflussung tatsächlich stattgefunden hat (34). „Sie sind als Gegebenheiten definiert, die ein Risiko dafür schaffen, dass professionelles Urteilsvermögen oder Handeln durch ein sekundäres Interesse unangemessen beeinflusst wird.“ (1) Sekundäre Interessen müssen nicht immer materieller Art sein, auch das Streben nach akademischer Anerkennung kommt infrage.

In einer vor einigen Monaten erschienenen Publikation, die eine Standortbestimmung der multimodalen Brustkrebs-Früherkennung darstellt, wurde über falsch positive Befunde und Überdiagnosen ausführlich berichtet, falsch negative Befunde wurden nicht erwähnt, das Wort Intervallkarzinom kam nicht vor (11). Der korrespondierende

rende Autor gab für sich und zehn Koautoren die Erklärung ab, dass es keinen Interessenkonflikt gebe, wobei ein Koautor einer der Protagonisten des Screenings in Deutschland ist. Die Frage ist nicht, ob er einen Interessenkonflikt hat, sondern warum behauptet wird, dass kein Interessenkonflikt besteht. Auf eine schriftliche Anfrage gab der korrespondierende Autor keine Antwort.

Karl Raimund Popper hat gesagt, dass es in der Naturwissenschaft keine Wahrheit gibt, nur optimal geprüfte Theorien, die durch Falsifikation verbessert werden müssen: Die These, dass alle Schwäne weiß sind, sollte nicht dadurch untermauert werden, dass man durchs Land fährt und weiße Schwäne zählt; findet sich der erste schwarze Schwan, ist die These widerlegt (28). Die als „Mammografie-Skandal im Ruhrgebiet“ beschriebenen Vorkommnisse um das Screening in Essen haben deutschlandweit große mediale Aufmerksamkeit hervorgerufen; sie lassen auf den ersten Blick an den schwarzen Schwan denken; die anhängigen Gerichtsverfahren werden Licht ins Dunkel bringen. Einer, der es jetzt schon weiß, ist der Leiter des Referenzzentrums Münster: „Die Berichte dürfen meines Erachtens aber nicht dazu führen, dass die guten methodischen Ansätze und die bisher erreichten Ergebnisse grundsätzlich infrage gestellt werden. (...) Es sollte allen Beteiligten klar sein, dass der Nutzen der Früherkennung durch eine Diagnosevorverlagerung nur mit einem Mammografie-Screening-Programm auf sehr hohem Qualitätsniveau erreichbar sein wird.“ (36)

Wie Kompetenz und Interessen professionell vermengt werden, zeigte sich auf dem 60. Kongress der Deutschen Gesellschaft für Gynäkologie und Geburtshilfe in der Hauptsitzung „Mammografie-Screening Deutschland: 9 Jahre – Was hat das gebracht?“ (4). Vorsitzende und Referenten wurden schon lange zuvor festgelegt – einigte Referenten der

letzten Jahre sind wahrscheinlich auch für die nächsten Kongresse bereits gesetzt. Kritikern des Screenings bleibt nur, sich als Diskutanten zu versuchen, was dadurch erschwert wird, dass Fragen zu den Referaten vorwiegend von den Vorsitzenden selbst gestellt werden. Leiter der Referenzzentren spielen sich die Bälle zu, belobigen ihre Vorträge gegenseitig, dogmatisieren ihre Sicht der Dinge und lassen zu, dass unliebsame Medienberichte gegeißelt und Kritikern rein monetäre Motive unterstellt werden – von Streitkultur keine Spur. Zwei Bemerkungen aus dem Auditorium wurden mit nichtssagenden Floskeln abgetan, eine in Aussicht gestellte abschließende Diskussion mit Rede und Gegenrede kam aus Zeitgründen nicht zustande.

### Zusammenfassung und Ausblick

- Screening-Untersuchungen liegen stets im Niedrigprävalenzbereich. Niedrige Prävalenzen ergeben auch bei sehr guter Screening-Qualität nur geringe positive prädiktive Werte, viele falsch positive Befunde, die Übertherapien auslösen, müssen zwangsläufig hingenommen werden. In allernächster Zeit könnte es aber gelingen, ein individuelles genetisches Risikoprofil zu erstellen (29). Genetisch stark belastete Frauen sind eher einem mittleren Prävalenzbereich zuzuordnen; diese Frauen profitieren dann wirklich von der Mammografie.
- Das Mammografie-Screening verfehlt in Deutschland mit der konstant niedrigen Teilnehmerrate (23) die Vorgaben der Europäischen Leitlinien bei Weitem. Dieses Faktum mit der Prämisse eines Paradigmenwechsels verstecken oder wegredden zu wollen, belegt zwar die Kreativität der Akteure, kann aber nicht darüber hinwegtäuschen, dass bei sehr geringer Teilnehmerzahl der vermutete Erfolg, bezogen auf die Brustkrebssterblichkeit in

Deutschland, nur sehr gering sein kann. Das deutsche Screening-Programm wird auch zukünftig die mit Bedacht gemachten Vorgaben der Europäischen Leitlinien klar verfehlen, es strebt diese ja schon gar nicht mehr an.

- Die „Stufe-3-Leitlinie Brustkrebsfrüherkennung in Deutschland“ von 2002 hat sich auf Frauen aller Altersgruppen bezogen und hat die gesamte Diagnosekette erfasst (sie war bereits auf dem Cover abgebildet): Gespräch, Risikoberatung und ärztliche Tastuntersuchung gehören ebenso dazu wie die Mammografie und die zusätzliche apparative Diagnostik mittels Ultraschall. Natürlich können, auch wenn das anders behauptet wird (4, 36), in einer Stufe-3-Leitlinie die Struktur- und die Prozessqualität ebenso gut wie im Screening umgesetzt werden. In einer Exit-Strategie ist darauf zu achten, dass die Expertise der Screening-Mammografie-Reader genutzt, verloren gegangene Strukturen reaktiviert und die patienten-nahen Möglichkeiten gut vernetzter Arztpraxen qualitätsgesichert eingebunden werden (27). Die zu erwartende Ergebnisqualität wird bei Anwendung der gesamten Diagnosekette besser sein, da dann auch mammo-graphisch okkulte Karzinome gefunden werden können.
- Interessenkonflikte müssen gerade beim Mammografie-Screening fest in den Blick genommen werden. Die Interessengemeinschaft Mammografie dominiert inzwischen nicht nur wissenschaftliche Kongresse, sondern macht sich zusehends auch im politischen Mainstream unentbehrlich. Die Nähe des von der Kooperationsgemeinschaft eingerichteten Referenzzentrums Mammografie in Münster zum Institut für Klinische Epidemiologie an der Universität Münster, dem die Studie zur Mortalitäts-evaluation in Deutschland

anvertraut wurde, fällt unangenehm auf. Auch in Berlin (Gigerenzer), Freiburg (Antes) und in Hamburg (Mühlhauser) gibt es gute Epidemiologen.

- Die Geschäftsleiterin der Kooperationsgemeinschaft geht, anders als die Deutsche Gesellschaft, davon aus, dass das Screening in Deutschland erst seit 2009 flächendeckend eingeführt war und dass sich der von ihr vermutete positive Screening-Effekt demzufolge erst 2019 bzw. 2021 zeigen wird (17). Der Deutsche Bundestag konstatiert, dass die altersstandardisierte Sterberate in den sieben Jahren vor Einführung des Mammografie-Screenings (1998–2005) in ähnlichem Umfang gesunken ist wie in den sieben Jahren seit Einführung (2006–2012), und erwartet mit Spannung die Ergebnisse der Machbarkeitsstudie aus Münster, die 2015 vorgelegt werden wird (5). Die Neubewertung des Screenings muss spätestens zu diesem Zeitpunkt erfolgen. Ein weiterer Aufschub wäre ein reines Spiel auf Zeit, das nur der Interessengemeinschaft zugutekommt.

Nicht akzeptabel wäre, die bisher vorgelegten geschätzten Zahlen 2015 durch neue geschätzte Zahlen – vielleicht auf der Basis einer weiteren Computersimulation – zu ersetzen. An der Stelle muss an Ingvar Andersson erinnert werden: Der Hauptuntersucher der Malmö-Studie sagte 1999, dass es noch einige Jahre dauern werde, bis der Einfluss auf die Mortalität in Schweden nachgewiesen werden könne; elf Jahre später konnte noch immer kein Einfluss nachgewiesen werden – aber niemand kümmernte sich mehr darum (13).

## Literatur

1. AWMF online: Empfehlungen der AWMF zum Umgang mit Interessenkonflikten bei Fachgesellschaften; abgefragt am 22.09.2014.
2. Biller-Andorno N, Jüni P: Abolishing mammography screening programs? A view from the Swiss Medical Board. *N Engl J Med.*; online publ. 16. April 2014; doi: 10.1056/NEJMp1401875
3. Bock K: Persönliche Mitteilung am Dt. Kongress für Gynäkologie und Geburtshilfe 2012.
4. Bock K, Heywang-Köbrunner SH: Mammografiescreening Deutschland: 9 Jahre – Was hat das gebracht? 60. Kongress der Deutschen Gesellschaft für Gynäkologie und Geburtshilfe, 8.–11.10.2014, München.
5. Deutscher Bundestag: Wissenschaftlicher Stand zu Nutzen und Risiken des Mammographie-Screenings. Drucksache 18/2299 vom 07.08.2014.
6. Dronkers JH, Hendriks J et al: Radiologische Mammadiagnostik, Thieme 1999.
7. European Commission: European guidelines for quality assurance in breast cancer screening and diagnosis, 2006.
8. Feiten S et al: Breast cancer morbidity – questionnaire survey of patients on the long term effects of disease and adjuvant therapy. *Dtsch Arztebl Int* 2014; 111: 537–44; doi: 10.3238/arztebl.2014.0537.
9. Fracheboud J, de Gelder R et al: Breast screening success – Dutch. *Wien*, 22. März 2012.
10. Gigerenzer G et al: Helping doctors and patients make sense of health statistics. *Psychol Sci Public Interest* 2008; 8(2): 53–96.
11. Golatta M et al: Multimodale Brustkrebs-Früherkennung – wo stehen wir heute?. *Frauenarzt* 55 (2014) Nr. 5, 438–45.
12. Gøtzsche PC, Olsen O: Is screening for breast cancer with mammography justifiable?. *Lancet* 2000; 355: 129–34.
13. Gøtzsche PC: Mammography Screening: Truth, Lies and Controversy. Radcliffe Publishing 2012.
14. Hanahan D, Folkman J: Patterns and Emerging Mechanisms of the Angiogenic Switch during Tumorigenesis. *Cell* 1996 Aug 9;86(3):353–64.
15. Heidinger O et al: The Incidence of Interval Cancers in the German Mammography Screening Program: Results from the Population-Based Cancer Registry in North Rhine-Westphalia. *Dtsch Arztebl Int* 2012; 109(46): 781–7; doi: 10.3238/arztebl.2012.0781.
16. Heinen-Kammerer T: Screening von 12.000 Frauen für ein gerettetes Leben. *Nsi/aerzteblatt.de* 2014.
17. Heinrich C: Persönliche Mitteilung vom 06.06.2014.
18. Hense HW: Die Sensitivität des Mammografie-Screening-Programms ist gut. [www.aerzteblatt.de/nachrichten/59496](http://www.aerzteblatt.de/nachrichten/59496)
19. Jaggar K: Breast Cancer Mammograms: overrated – and over-diagnosing Women. *The Guardian*. Posted on April 28, 2014.
20. Jørgensen KJ, Zahl PH, Gøtzsche PC: Breast cancer mortality in organized mammography screening in Denmark: comparative study. *BMJ* 2010; 340: c1241
21. Katzenmeier C: Mammographie-Screening: Rechtsfragen weitgehend ungeklärt. *Dtsch Arztebl* 2006; 103(16): A-1054 / B-892 / C-861
22. Kaulen H: Hoffnungslos übertherapiert? *Frankfurter Allgemeine Zeitung*, Nr. 100, 1. Natur und Wissenschaft.
23. Kooperationsgemeinschaft Mammographie: Evaluationsbericht 2010, Ergebnisse des Mammographie-Screening-Programms in Deutschland, Februar 2014.
24. Kolb TM, Lichy J, Newhouse JH: Comparison of the performance of screening mammography, physical examination and breast US and evaluation of factors that influence them: an analysis of 27825 patient evaluation. *Radiology* 225 (2002): 165–175.
25. Kürzl R: Grundlagen diagnostischer Tests und Screeningverfahren. In: Schneider H, Husslein WP, Schneider KTM: Die Geburtshilfe, Springer 2006.
26. Placzek U et al: Sind Mammographie-Screening und Risikomanagement kompatibel? *Frauenarzt* 47 (2006) Nr. 10, 898–903.
27. Placzek U: Mammografie-Screening: ineffektiv oder zukunftsfähig? *Frauenarzt* 55 (2014) Nr. 6, 540–544.
28. Popper KR: Logik der Forschung. Springer, Wien 1935, Hrsg. Herbert Keuth, Mohr Siebeck, Tübingen 2005 (11. Aufl., ISBN 3-16-146234-3).
29. Schmutzler R et al: Beratung, Genetische Testung und Prävention von Frauen mit einer familiären Belastung für das Mamma- und Ovarialkarzinom. Interdisziplinäre Empfehlungen „Familiärer Brust- und Eierstockkrebs“ der Deutschen Krebshilfe. *Medgen* 15 (2003).
30. Stähler E et al: Angiogenetic switch. *Der Privatarzt Gynäkologie* 5 (2012), 26–32.
31. Mukhtar TK, Yeates D, Goldacre MJ: Breast cancer mortality trends in England and the assessment of the effectiveness of mammography screening: population-based study. *J R Soc Med Jun* 2013; 106(6): 234–242.
32. Weedon-Fekjær H: Beobachtungsstudie Trondheim. *BMJ* 2014; doi: 10.1136/bmj.g3701.
33. Weymayr C: Dokumentation 2010 in Auftrag der Kooperationsgemeinschaft Mammographie.
34. Wikipedia: Interessenkonflikt; abgefragt am 22.09.2014.
35. [www.cochrane.dk](http://www.cochrane.dk). 2008.
36. Zylka-Menhorn V: Mammographie-Skandal im Ruhrgebiet. Privatfehde oder Versagen der Kontrollorgane? *Dtsch Arztebl* 2014; 111(23–24), A-1098.

## Interessenkonflikt

Der Autor erklärt, dass keine Interessenkonflikte bestehen.



## Autor

**Dr. Ulrich Placzek**  
Straubinger Str. 30c  
80687 München  
uplaczek@web.de  
Tel. 0170 4841463