

Johannes Schmidt

Die Brustkrebs-Vorsorgeuntersuchung

Die kritische und praktisch relevante Beurteilung wissenschaftlicher Daten¹

Das Problem der Brustkrebs-Vorsorgeuntersuchung wird allzu häufig auf die Frage beschränkt: Kann die Früherkennung tatsächlich die Sterberate des Brustkrebses (Mammakarzinoms) verringern oder nicht? Die Faszination an der klinisch interessanten Frage, ob denn nun durch eine frühe Entfernung von Krebsgewebe der Krebs besiegt werden kann oder nicht, verlegt oft den Blick für eine umfassendere, und für die Praxis eigentlich entscheidende Sichtweise. Von einer klinisch-epidemiologischen Betrachtungsweise her muss die Frage denn vielmehr lauten: Wie sieht das Verhältnis aus zwischen einerseits der Verringerung der Sterberate und andererseits den durch die Früherkennung möglichen negativen gesundheitlichen und sozialen Auswirkungen?² Mit anderen Worten: Überwiegt das durch die Früherkennung verminderbare Leiden tatsächlich das dabei erzeugte Leiden?

»Natural History«,³ des Mammakarzinoms

Historisch hat sich die Vorstellung entwickelt und festgesetzt, das Mammakarzinom folge einem strengen Wachstums- und Verbreitungsmuster, welches durch ein zunächst lokales Wachstum im Brustdrüsenbereich mit einer anscheinenden Ausbreitung über die Lymphwege in die axillären Lymphknoten gekennzeichnet ist. Diese lymphogene Ausbreitung erfolgt erst einige Monate bis einige Jahre nach einem streng auf die Brust beschränkten Krebswachstum. Erst dann, gemäss dieser Vorstellung, verbreitet sich das Krebsgewebe hämatogen (über die Blutwege) in den ganzen Körper, um verstreut Metastasen zu bilden (Henderson 1980). Die Bedeutung einer Früherkennung des Mammakarzinoms (im vermeintlich noch streng lokalen Stadium) stützt sich sehr stark auf diese Vorstellung ab.

Auch heute noch ist die »natural history« des Mammakarzinoms in weiten Teilen unerforscht. Verschiedene Hypothesen über Wachstum und Ausbreitung des Brustkrebses müssen in Betracht gezogen werden. Neben der erwähnten herkömmlichen Theorie eines geordneten lokalen Wachstums mit erst sekundär stattfindender Ausbreitung von Metastasen, muss aufgrund verschiedener Beobachtungen angenommen werden,

dass der Brustkrebs sich sehr häufig wie ein »systemischer« Krebs verhält, bei welchem praktisch von Anfang an ein disseminiertes (im ganzen Körper gestreutes) Wachstum von Krebszellen und -gewebe stattfindet. Der Tumor in der Brust ist dabei nur das auffälligste Symptom eines eigentlich im ganzen Körper stattfindenden Krankheitsprozesses. In einer histo-pathologischen Untersuchung zur Abklärung dieser Frage konnten tatsächlich nur schon im Knochenmark von 28 % der Patientinnen mit Brustkrebs im Frühstadium (»minimal cancer«: in situ oder ≥ 1 cm) mikroskopische Metastasen nachgewiesen werden (Redding 1983).

Es ist deshalb gut möglich, dass es zwei grundsätzlich verschiedene, histologisch nicht unterscheidbare, (Verlaufs-) Formen von Brustkrebs gibt mit entweder streng lokalem Wachstum oder primärer Dissemination. Die offene Frage ist, wie gross deren relative Häufigkeit sein könnte und ob sich Unterscheidungsmerkmale finden liessen (Henderson 1980).

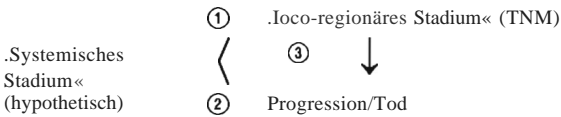
Deshalb ist es ein Irrtum anzunehmen, dass die zunehmend schlechtere Prognose (Ueberlebenschancen) mit fortschreitendem (TNM-) Stadium einen tatsächlichen Krankheitsverlauf von Stadium I über Stadium II und III bis zum Stadium IV und zum Tod *beweisen* würde. Die gebräuchliche klinisch-histologische Stadien-Einteilung (TNM-Staging) 'beruht auf dem überholten Modell streng lokalen Wachstums und diese Einteilung wurde vor allem entwickelt, um stadiengerechte Behandlungsmodalitäten und diejenigen Patientinnen zu bestimmen, bei welchen das Mammakarzinom noch durch eine radikale Mastektomie »kurierbar« sei. Tatsächlich ist dies aber unmöglich. Denn Patientinnen mit einem früh entdeckten Karzinom, welche einer Halsted'schen radikalen Mastektomie unterzogen werden, weisen weit häufiger ferne Rezidive auf als im Bereich der wegoperierten Brust oder der »ausgeräumten« Axilla (Henderson 1980). Obwohl das »Staging« (Stadien-Zuordnung) prognostische Bedeutung hat, repräsentiert es also nicht unbedingt den natürlichen Verlauf der Krankheit oder nur in einer bisher nicht genauer definierbaren Untergruppe von Mammakarzinomen.

Ein alternatives Modell der natural history des Brustkrebses mit besserer Erklärungskraft kann entworfen werden, welche diesen Beobachtungen gerecht wird. Die enge Beziehung zwischen TNM-Stadium und Prognose beweist nämlich noch nicht, daß die bekannten klinischen und histologischen Merkmale der entsprechenden Stadien wirklich die Ursache der weiteren Krankheitsentwicklung bis hin zum Tod darstellen. Falls die Krankheit *primär systemisch* verläuft, können die her-

kömmlichen TNM-Stadien zwar einen fortschreitenden Krankheitsverlauf *anzeigen*, ein hypothetisches »systemisches Stadium« könnte aber den Faktor ausmachen, welcher den Krankheitsverlauf und die Prognose *bestimmt*. Die Beziehung zwischen dem lokal definierten TNM-Stadium und der Prognose könnte also durch ein »systemisches Stadium« begründet (konfundiert) sein; dem TNM-Stadium käme lediglich der Charakter eines Epiphänomens (nur *Indikator* aber nicht *Verlaufs-Faktor*) zu. Dies ist in *Abbildung 1* dargestellt:

Abb.1:

Modell eines disseminierten Wachstums des Mammakarzinoms



I: Konfundierende Assoziation; 2: Kausale (.unabhängige) Assoziation, 3: Unechte (.abhängige) Assoziation / Kausal nur in Fällen streng lokaler Determination des Verlaufs

Wie aus dem Modell ersichtlich, ist gar nicht unbedingt zu erwarten, daß eine frühzeitige »lokale« Erkennung eines Mammakarzinoms und dessen (radikale) Entfernung die Progression der Krankheit verhindert, denn die möglicherweise entscheidende **pathogenetische** Beziehung zwischen dem »systemischen Stadium« und der Progression (in der Abb. Pfeil 2) wird bei einer lokalen Manipulation nicht berührt. Es wird auch klar, daß die Verhütung eines Stadium II- oder III-Karzinoms aus einem minimal cancer nicht unbedingt bedeutet, daß die Krankheit zwangsläufig in ihrem natürlichen Verlauf gestoppt worden ist.

Es gibt keine Hinweise, daß medizinische Maßnahmen einen eindeutigen Effekt auf die Brustkrebs-Sterberate gehabt hätten. Veränderungen in Inzidenz und in Überlebensraten sind zwar beobachtet worden, können aber einfach Ausdruck medizinischer Vorsorge-Aktivitäten sein, ohne daß sich am Verlauf der Krankheit etwas verändern muss. Betrachtet man die absolute Brustkrebs-Sterberate als das einzig unverfälschte Maß, so lassen sich seit über 30 Jahren keine Behandlungsfortschritte feststellen (Bailar 1986).

Der Nutzen der (radikalen) Mastektomie ist nie in geeigneten kontrollierten Studien festgelegt oder bewiesen worden, obschon diese »Therapie« auf (mindestens zum Teil) falschen Vorstellungen über Biologie und Wachstum des Mammakarzinoms beruht. Die von Halsted, dem Begründer der radikalen Brustamputation, behandelten Patien-

tinnen zeigten effektiv keinen eindeutig besseren Verlauf als etwa eine vergleichbare Gruppe unbehandelter Patientinnen des Middlesex Hospitals (Henderson 1980). Da keine vergleichende randomisierte Studien zum Nutzen der Brust-Operation (wie radikal auch immer) bei beschwerdefreien Karzinomen vorliegen, wissen wir effektiv nicht, ob und wann eine Operation den Krebstod verhindern kann. Wenn operiert wird, so zeigen immerhin mehr und mehr neue Studien, daß modifizierte brusterhaltende Operationsverfahren dem historischen radikalen Verfahren ebenbürtig sind, d.h. das Mammakarzinom genauso viel oder wenig heilen können, wenn auch solche Studien erst bei kleinen Krebsen gewagt worden sind (Henderson 1980).

Brustkrebs-Vorsorgeuntersuchungen wurden also eingeführt, ohne daß überhaupt Klarheit über eine wirksame *minimale* Therapie bestand.

Fallgruben klinischer Beobachtungen

und die Notwendigkeit vergleichender randomisierter Studien

Den in der klinischen Denkweise geformten Medizineren fällt es offensichtlich schwer, die Trugschluss-Möglichkeiten in diesen herkömmlichen Vorstellungen zu erkennen. Dies liegt wohl in einer Unkenntnis möglicher Verfälschungseffekte (»bias«), die offensichtlich einleuchtende klinische Beobachtungen begleiten. Wir können unbestreitbar beobachten, daß Patientinnen, bei welchen ein Mammakarzinom (durch Vorsorge-Untersuchungen) früh entdeckt wird, eine bessere Überlebenschance haben als diejenigen, bei welchen ein Karzinom erst in einem späteren Stadium festgestellt wird« Dennoch kann daraus nicht ein Nutzen einer Frühintervention abgeleitet werden, denn verschiedene Verfälschungseffekte machen es möglich, daß die Überlebenszeit früh erfaßter Karzinome immer weit günstiger erscheint, auch wenn sich am eigentlichen Verlauf gar nichts änderte (Sackett 1985). Es geht um den Latenzzeit-Verfälschungseffekt ("lead-time bias«), den Wachstumsgeschwindigkeit-Verfälschungseffekt (»length bias«), den Selektionseffekt von Gesünderen (»healthy screenee effect«) und um den Beurteiler-Verfälschungseffekt (»observer bias«).

Lead-Time Bias: Eine Früh-Diagnose verlängert die Überlebenszeit zwischen Diagnose und Tod zwangsläufig, einfach weil der »Krankheitsausbruch« mit der Frühdiagnose vorverlegt wird. Wenn also die Überlebenszeit des Mammakarzinoms durch die Vorsorgeuntersuchungen gestiegen ist, so *könnte* dies einfach bedeuten, daß die klinische Phase (Leidenszeit) länger geworden ist bei insgesamt gleich gebliebener Lebenserwartung.

Length Bias: Die Wachstumsgeschwindigkeit des Mammakarzinoms weist eine große Streubreite auf. Schnell wachsende Karzinome können oft durch die Vorsorgeuntersuchung nicht erfaßt werden, weil sie beim ersten Screening noch zu klein waren, um erfaßt zu werden, aber schon vor dem nächsten Screening zu klinisch erfassbaren Tumoren ausgewachsen (»Intervall-Karzinome,,). Man schätzt, daß mindestens alle 4 Monate eine Mammographie durchgeführt werden müßte, um alle Karzinome im Frühstadium erfassen zu können (Haagensen 1981). Bis zu einem gewissen Grade sind deshalb screening-entdeckte Karzinome eine *prognostisch gutartige Auswahl* langsam wachsender Krebse.

Healthy Screenee Effect: Fast regelmäßig sind Personen, welche an Vorsorge-Programmen teilnehmen, von vornherein gesünder als die »Verweigerer". So wiesen in der Health Insurance Plan-Studie (HIP) in New York die Screening-Teilnehmerinnen in Bezug auf die Gesamt-Mortalität eine viel bessere Prognose auf als die »Verweigerinnen« (Shapiro 1977, Sackett 1985):

Screening-Teilnahme und Todesraten - HIP Trial

	Gesamtslerberale pro 10000 Frauenjahre
Screening-Gruppe: Teilnehmerinnen	42
Verweigerinnen	77
Vergleichsgruppe	54

Screening-entdeckte Karzinome sind somit bei einer *Auswahl* von Frauen zu finden, die allgemein eine bessere Lebenserwartung haben.

Observer bias: Es ist möglich, daß fragliche Atypien im Rahmen eines Screening-Programmes als Mammakarzinome überdiagnostiziert werden aus dem Bestreben heraus, kein Karzinom zu verpassen. Screening-entdeckte Karzinome können deshalb eine *Auswahl* gutartiger »Krebse« sein, insofern falsch diagnostizierte Krebse gutartige Nicht-Karzinome sind. (Editorial Lancet 1981). In einem grossen Screening-Programm (BCDDP) war es bei 17 % der kleinen Tumoren oder bei 6 % aller entdeckten »Brustkrebse« in einer Nachbeurteilung tatsächlich strittig, ob es sich wirklich um Karzinome handelte (Working Group 1977).

Die möglichen Verfälschungs-Effekte können nur in relativ aufwendigen *vergleichend* randomisierten Studien⁴ ausgeschaltet werden (Sackett! 1985). Obwohl Fallkontroll-Studien⁵ wegen des kleineren Aufwandes attraktiv erscheinen, können auch diese nicht alle Verfälschungen ver-

meiden. Length bias und healthy screenee effect kommen auch in diesem Studien-Design zum Zug. Das relative Risiko⁶ von an Vorsorgeuntersuchungen teilnehmenden Frauen in solchen Studien wird durchwegs stark unterschätzt (Verbeek 1984, Coilelte 1984), was die innewohnenden Verfälschungs-Effekte aufzeigt.

Da es letztlich darum **gehen** muß, alle positiven und negativen Effekte eines Programmes in ihrer (quantitativen) Bedeutung abzuschätzen, um zu beurteilen, ob ein Screening mehr Leiden verhindert als schafft, müssen angebliche Nutzensbeweise aus Fallkontroll-Studien als wenig relevant zurückgewiesen werden. Immer noch, schon fast ein Skandal, erscheinen Publikationen in angesehenen Zeitschriften, welche aufgrund von Beobachtungs-Studien⁷ den Nutzen des Brustkrebs-Screenings belegen wollen (Rodes 1986).

Kontrollierte randomisierte Studien

Bisher liegen Daten von zwei randomisierten Screening-Studien vor: aus New York (HIP) und Schweden. Weitere Studien dieser Art werden zur Zeit in Kanada (Miller 1981) und Edinburgh (Roberts 1984) durchgeführt.

Eine moderne Mammographie, wie in der neueren Schwedischen Studie angewandt, ist weit sensibler und mit einer geringeren Strahlenbelastung behaftet als die Technik, welche für die ältere HIP-Studie verwendet werden mußte. In der großen (nicht vergleichenden) BCDDP-Studie in den USA (Behrns 1979) konnte gezeigt werden, daß bei Verwendung einer modernen Technik ein gutes Drittel der Mammakarzinome schon im Stadium I (lokalisiert, nicht grösser als 2 cm) entdeckt werden **kann** und daß ein Drittel dieser kleinen Krebse bei Frauen unter 50 Jahren zu finden war. Obschon in der älteren HIP-Studie die manuelle Brustuntersuchung noch eine gewisse Rolle spielte, konnte nun gezeigt werden, daß die moderne Mammographie weit sensibler ist, und daß Selbstuntersuchung und sogar die Palpation durch einen erfahrenen Arzt für die Entdeckung der kleinen Thmoren keinen Wert haben (Hendricks 1985).

Die Schwedische Studie ist die jüngste, von welcher Ergebnisse vorliegen, und sie repräsentiert, was mit einem modernen Mammographie-Programm möglich ist. Über die Hälfte der Mammakarzinome in der Screening-Gruppe waren kleine Karzinome im Stadium I, jedoch nur 35 % in der Kontroll-Gruppe ohne Mammographie-Screening (Tabar 1985a) ($p \ll .001$, vgl. *Tabelle 3*).

Die Studie wurde an der gesamten weiblichen Bevölkerung der beiden Grafschaften Koppaberg und Oestergötland durchgeführt. In

Kopparberg wurden mehr als 70000 Frauen zwischen 40 und 74 Jahren in die Studie einbezogen, 2/3 wurden nach dem Los der Screening-Gruppe zugeteilt, 1/3 der Kontroll-Gruppe. Die Randomisierung fand auf Gemeinde-Ebene statt und wurde nicht individuell durchgeführt. Über die verschiedenen Screening- und Abklärungs-Aktivitäten, liegen Daten aus Kopparberg für die Altersgruppe zwischen 40 und 69 Jahren vor (Tabár 1985a). Mammographien (I-Bild-Schrägaufnahmen) wurden als einzige Screening-Methode verwendet und fanden in der jüngeren Altersgruppe alle 24 Monate, in der älteren alle 33 Monate statt. Die Teilnahme-Rate war mit rund 90 % außerordentlich gut. Zumindestens in einer Gesellschaft wie in Schweden sind Massen-Untersuchungen dieser Art somit praktisch durchführbar.

Screening-Aufwand

Nach 7 Jahren waren (bei den Teilnehmerinnen, 40-69 Jahre) in der Screening-Gruppe die folgenden Abklärungen vorgenommen worden:

Tabelle 1:

Brustkrebs-Vorsorgeuntersuchung

7-Jahresaufwand pro 10000 Frauen (40-69 Jahre, n=33641) - Kopparberg Studie

Einfach-Mammographien (I-Bild)	20100
Abklärungs-Mammographien mit mehreren Projektionen (in einer 2. Sitzung)	790
Nadel-Biopsien/klinische Untersuchungen	270
Schnellschnitt-Histologien	140
Früh-entdeckte Karzinome:	
Früh-Maslektomien/-Lumpektomien	80

Pro früh-entdecktem Karzinom waren somit rund 250 einfache und 10 komplette Mammographien, sowie 3,3 Nadelbiopsien und 1,7 Schnellschnitt-Histologien notwendig. Die Gesamtzahl der in 7 Jahren aufgetretenen und entdeckten Mammakarzinome ist in Tabelle 2 wiedergegeben:

Tabelle 2: Kopparberg-Studie

7-Jahres-Inzidenz von Mammakarzinomen in Screening-Gruppe (40-69 Jahre)

Diagnosezeitpunkt	Anzahl
Randomisierung bis Screening-Beginn	12
Screening	321
Interval-Karzinome	105
Bei Verweigerinnen	27
Total (n = 33461)	465

Aus diesen Zahlen lassen sich *Sensibilität* und *Spezifität* des Mammographie-Screenings ableiten (vgl. Abholz, in diesem Band).

Die errechnete *Sensibilität* beträgt 72 % für die Teilnehmerinnen (»on-screening«8), resp. 67 % für die ganze Zielgruppe (»intention-to-screen«9). Zwei Drittel (67 %) aller Mammakarzinome ließen sich durch das Vorsorge-Programm, welches eine hohe Teilnehmer-Rate aufwies, demnach früher erkennen (und behandeln). Ein Drittel konnte auch mit dem Screening nicht früher erfaßt werden. Der errechnete *Negative Prädiktive Wert* beträgt 99.6 %. Nimmt man am Screening teil, so weiß man bei durchgehend negativen Befunden somit mit 99.6 % Wahrscheinlichkeit, daß kein Karzinom auftreten wird. Die Wahrscheinlichkeit, daß in 7 Jahren *kein* Karzinom auftritt oder gefunden wird, ist jedoch von vornherein schon 98.6 %, also ebenfalls sehr groß. Ein negatives Manunogramm hat für eine teilnehmende Frau somit kaum einen praktischen Informations- oder Beruhigungs-Wert, da für ein Individuum eine Chance von 98.6 % oder 99.6 % als identisch empfunden werden dürfte.

Die *Spezifität* der Mammographie erscheint mit errechneten 92 % recht hoch. Es ist jedoch ein häufiger Fehler, zu glauben, daß eine hohe Spezifität einer Untersuchung bedeute, ein positiver verdächtiger Befund sei mit hoher *Wahrscheinlichkeit* tatsächlich positiv. Diese Wahrscheinlichkeit wird vielmehr durch den *Positiv Prädiktiven Wert*¹⁰ ausgedrückt, welcher bei den Inzidenz-Verhältnissen in Koppberg lediglich 10 % für die Einfach-Mammographie beträgt und 30 % für die komplette Mammographie, die an einem späteren Untersuchungstag stattfinden muß. Eine zunächst verdächtige erste Mammographie bedeutet also nur in einem Zehntel der Fälle, und in einem Viertel bis Drittel positiver kompletter Mammographien, daß tatsächlich ein Karzinom vorliegt. 90 % werden zunächst umsonst einem Verdacht und weiteren Abklärungen ausgesetzt (das Ausmaß der dadurch erzeugten psychischen Belastungen ist bisher kaum untersucht und dokumentiert worden). Dieser Aufwand und die dabei möglichen Nebenwirkungen können nur im Vergleich mit den nützlichen Effekten bewertet werden.

Screening-Ergebnisse

Die Studien-Autoren legen eine »signifikante Reduktion des fortgeschrittenen Mammakarzinoms« im Titel und in den Schlussfolgerungen ihrer Publikation (Tabar 1985a) nahe. Die entsprechenden Resultate sind in *Tabelle 3* zusammengefasst:

Tabelle 3:

Die Koppa-berg-Studie. Anzahl und **Verteilung** der Mammakarzinome in verschiedenen **Stadien** (**UICC TNM-Klassifikation 1978**) in **Screening- und Kontroll-Gruppe**

	bis Stadium I	op-NX.	Stadium II-IV	Total	mit Metastasen
Screening-Gruppe (n = 33461)	253 (54%)	49	163 (35%)	465	117 (25%)
Kontroll-Gruppe (n = 16359)	62 (34%)	4	116 (64%)	182	68 (37%)

op-NX. =infiltrierend, (20 mm Durchmesser, Lymphknoten-Histologie fehlend, deshalb potentiell unter-staged. In Klammern: Anteil in % — • $p < .001$ (Chi-Quadrat =43,8) - •• $p = .002$ (Chi-Quadrat =9,5)

Tabelle 3 spricht tatsächlich eine deutliche Sprache, was die statistisch hochsignifikante Unterschiede zwischen Screening- und Kontroll-Gruppe in der proportionalen Verteilung von Früh- und Spät-Stadien betrifft. Der Anteil Karzinome in fortgeschrittenen Stadien ist in der Kontrollgruppe ohne Screening fast doppelt so hoch. Ein erstes Problem mit der Qualität dieser Daten besteht allerdings darin, daß Angaben fehlen, ob das Staging der einzelnen Patientinnen blind erfolgte, d.h. in Unkenntnis des Beurteilers über die Gruppenzugehörigkeit. Verfälschungseffekte einer nicht blinden Beurteilungweise (observer bias) können erheblich sein und entstehen, weil Beurteiler meist immer, wenn auch ganz subtil, voreingenommen sind.

Wenn wir uns nun an die obigen Ausführungen erinnern, so müssen wir aber feststellen, daß uns die proportionale Verteilung verschiedener Krebsstadien eigentlich nichts aussagen kann über eine mögliche Verhütung fortgeschrittener Krebsstadien. Dies wird deutlich, wenn wir uns statt der Proportionen die absoluten Inzidenzen¹¹ vor Augen halten:

Tabelle 4:

Inzidenz verschiedener Brustkrebs-Stadien in Screening- und Kontroll-Gruppe (7-Jahres-Inzidenz, pro 1000 Frauen). Koppa-berg Studie.

	bis Stadium I	op-NX. + Stadium II-IV	Stadium U-IV	mit Metastasen	Total
Screening-Gruppe	7,6	6,3	4,9	3,5	13,9
	+	n.S.*	++	n.S.**	+++
Kontroll-Gruppe	3,8	7,3	7,1	4,2	11,1
	+ $p < .001$	• $p = .20$	+ + $p = .002$	" $p = .26$	+ + + $p = .010$

Wir sehen nun, daß bei einem Früherkennungs-Programm der proportionale Anteil früher Stadien an allen Karzinomen logischerweise größer wird, und damit der Anteil fortgeschrittener Stadien geringer, auch wenn sich an der Inzidenz fortgeschrittener Stadien vielleicht gar nichts ändert, denn es werden ja auch deutlich mehr (vor allem frühe) Krebse erfaßt. Tatsächlich sind aus *Tabelle 4* keine eindeutigen Inzidenz-Unterschiede fortgeschrittener Karzinome festzustellen, wenn auch die alleinige Betrachtung des Inzidenz-Unterschied bezüglich Stadium II-IV den Schluß nahelegen könnte, durch die Vorsorge sei tatsächlich eine Verhinderung fortgeschrittener Stadien möglich.¹²

Dies zeigt, wie vorsichtig man immer wieder sein muß im Akzeptieren "wissenschaftlicher« Schlussfolgerungen. Mit der gleichen (falschen) Logik könnte man nämlich auch folgern, die Krebsvorsorge verursache Krebse, weil ja in den 7 Jahren in der Screening-Gruppe mehr klinisch erkannte Karzinome aufgetreten sind (der 20 %-ige Unterschied ist statistisch signifikant).

Mortalitäts-Ergebnisse

Nach allen bisherigen Ausführungen und Mahnungen vor Studien-Verfälschungseffekten und Fehlinterpretationen ist es nun natürlich sehr aufschlußreich, Mortalitäts-Ergebnisse dieser kontrollierten randomisierten Studie aus Schweden zu sichten. Erste 7-Jahres-Ergebnisse diesbezüglich sind für beide Grafschaften veröffentlicht worden (Tabar 1985b):

Tabelle 5:

Brustkrebs-Mortalität in der Schwedischen Screening-Studie nach durchschnittlich 7 Jahren Verlaufszeit. Frauen zwischen 40 und 74 Jahren.

SIUdienpopulation	Brustkrebs Todesfälle	Inzidenz pro 10000	Inzidenz-Unterschied pro 7 Jahre	relatives Risiko"
Screening-Gruppe (n = 39051) <i>Kopparberg</i>	51	13,1	7,6 (.2,15.0)	.63+
Kontroll-Gruppe (n = 18846)	39	20,7		
Screening-Gruppe (n = 39034) <i>Oestergoetland</i>	36	9,2	3,2 (-1.5,7.8)	_.74++
Kontroll-Gruppe (n = 37936)	47	12,4		
Screening-Gruppe (n = 78085) <i>Beide kombiniert</i>	87	11,1	4,0 (.04,8.0)	.69+++
Kontroll-Gruppe (n = 56782)	86	15,1		

* wiedergibt Mortalitätsrisiko der Screening- im Vergleich zur Kontroll-Gruppe; RR/DR von .69 bedeutet Mortalitäts-Senkung um 31 %. Der Inzidenz-Unterschied (individual attributable benefit) bezeichnet die Wahrscheinlichkeit, mit welcher ein Individuum vom Screening in Form eines vermeintlichen Todesfalles profitieren kann. In Klammern: 95% Vertrauensintervalle. / + p = .03; ++ / p = .01 (Mantel-Haenszel Odds Ratio)

Ein Problem der in *Tabelle 5* vorliegenden Sterbe-Ziffern ist die fehlende Alters-Standardisierung, die in der Publikation nicht vorgenommen wurde. Die Altersstandardisierung ist in der randomisierten Studie zwar nicht in jedem Fall erforderlich, weil der Randomisierungsprozess ja eine gute Vergleichbarkeit schafft. Beim Gemeinde- und Stadtteil-bezogenen Randomisierungsverfahren dieser Studie können jedoch Verfälschungseffekte durch das Alter eine Rolle spielen. Jedenfalls bedürften die relativ großen Inzidenz-Unterschiede zwischen den beiden Graf-schaften (für **Kontroll-Gruppen**: $p = .02$) einer Erklärung, die in der Pu-blikation fehlt. Neben möglichen (undurchsichtigen) Unterschieden in der Studiendauer, könnten Alterseffekte dabei eine Rolle spielen.

Trotz dieser etwas mangelhaften Daten-Widergabe zeigt die Schwedi-sche Studie wohl, daß sich durch das Screening die Mammakarzinom-Mortalität um rund 30 % senken läßt. Obwohl die Ergebnisse in Oester-götland nicht signifikant sind (die Zufallsmöglichkeit ist mit einem p-Wert von 18 % groß), zeigt die kombinierte Gesamtstudie ein eindeutig signifikantes Resultat, welches nicht dem Zufall zugeschrieben werden kann. Die nun (mit den geäußerten Vorbehalten) bewiesene Reduzier-barkeit der Krebssterblichkeit durch die Früherkennung hat Signifikanz für die Theorie der natural history des Mammakarzinoms: *Mindestens* eine Untergruppe des Mammakarzinoms wächst also so streng lokal, daß eine Früherkennung mit früher lokaler Therapie das Karzinom zum Verschwinden bringen kann. Doch ist die *statistisch signifikante* Senkung der Todesrate auch »*praktisch signifikant*«? Diese Frage ist tat-sächlich erlaubt und relevant, wie wir sehen werden.

Eine aussagekräftige Nutzens-Beurteilung

Die prozentuale Mortalitäts-Reduktion sagt über die praktische Nütz-lichkeit einer Intervention kaum etwas aus. Die Beurteilung eines Therapie- und Interventionsnutzens anhand des Relativen Risikos (= der prozentualen Mortalitäts-Senkung) ist eigentlich als falsch zu betrachten. In *Tabelle 5* ist neben dem Relativen Risiko auch die absolute Risikoverminderung aufgeführt. Die 31 %-ige Reduktion der Mam-makarzinom-Sterblichkeit entspricht, mit diesem Parameter gemessen, 1 verhinderten Todesfal-ec-lunter 2500 Frauen im Zeitraum von 7 Jah-ren«. Das 95 %-Vertrauensintervall reicht von 1 verhinderten Todesfall unter 1250 Frauen bis zu unter 250000 Frauen! Wir sind also zu 95 % sicher, daß sich in 10 Jahren ein Todes-Ereignis bei rund 0 bis II von 10000 Frauen verhindern läßt, die zu 90 % regelmäßig am Screening teilnehmen. Wie diese Darstellung nahelegt, kommt es bei der Beur-

teilung der praktischen Nützlichkeit des Screenings also kaum darauf an, um wieviel Prozent sich die Sterberate senken läßt oder wieviele Aktivitäten zur Früh-Entdeckung eines Falles notwendig sind. Entscheidend sind vielmehr der absolute Inzidenz-Unterschied und Ausmaß notwendiger Intervention *pro verhindertem Todesfall*. Erst so lassen sich die gesamten nützlichen und unerwünschten Auswirkungen effektiv gegeneinander abschätzen.

Technischer Aufwand für die Verhütung eines Brustkrebs-Todesfalls

Während *Tabelle 1* und *2* über die »technische Güte« des Screenings Auskunft gegeben haben, zeigt *Tabelle 6* nun die viel entscheidendere gesundheitliche Bilanz der Vorsorge:

Tabelle 6:

Screening-Aktivitäten* pro **verhinderem** Krebsstod. Die Schwedische Studie.

Screening-Aktivität	pro 100 Frauen	pro verhindertem BR'Krebs-Todesfalli
I-Bild Mammographie	240	5900 (3000-580000) +
Komplette Mammographie	9,3	230 (120-23000)
Nadelbiopsie/Palpation	3,1	79 (40-7700)
Operative Biopsie	1,6	41 (20-4000)
Früh-Mastektomie/-Lumpektomie	1,0	24 (12-2300)

• Hochgerechnet von Koppberg-Studie. Weil in der entsprechenden Publikation (**Tabár** 1985a) lediglich die Screening-Aktivitäten aufgeführt sind, die zur Frühdiagnose von 273 von aber insgesamt 321 Karzinomen geführt hatten, und weil der Bericht sich auf nur 33641 Frauen (40-69 Jahre), die Mortalitätsangaben (**Tabár** 1985b) sich aber auf 39051 Frauen (40-74 Jahre) beziehen, wurde ein Faktor 1.36 ($321 : 273 \times 39051 : 33641$) verwendet, um die Screening-Aktivitäten für die ganze Gruppe, für welche Mortalitätsdaten vorhanden sind, zu schätzen. Zahlen gerundet. — **+** =95%-Vertrauensintervalle (mit95% Sicherheit liegt die wirkliche Anzahl irgendwo im angegebenen Bereich)

Die medizinische Profession weist immer wieder auf die 30 % Reduktion der Krebs-Sterblichkeit durch das Screening hin, um die Notwendigkeit der medizinischen Vorsorge und die entsprechende Lehrmeinung zu begründen. Doch wer hätte sich je überlegt, daß zur Verhinderung eines Mammakarzinom-Todes an die 10000 Mammographien notwendig sind (vgl. *Tabelle 6*), oder daß in 10 Jahren nur 1 von etwa 1750 Frauen (=4 von 10000 in 7 Jahren; *Tabelle 5*) direkt von einer Todesfall-Verhütung profitieren kann!?

Finanzieller Aufwand für die Verhütung eines Brustkrebs-Todes

Aufgrund von US-Daten lassen sich etwa finanzielle Kosten von 20 Dollar pro Frau und Jahr (Wertheimer 1986) berechnen. Dies ergibt aufgrund der Schwedischen Ergebnisse rund 350000 \$ pro verhütetem

Brustkrebs-Todesfall (Die Schwedischen Autoren veranschlagen geringere Kosten, die aber nicht generalisierbar erscheinen). Wie *Tabelle 7* zeigt, sind diese finanziellen Kosten dennoch vergleichsweise (Himmelstein 1984) gering.

Tabelle 7:

Brustkrebs-Screening: Kostenvergleiche

Intervention	verhindertes Ereignis	Kosten pro verhindertes Ereignis
Massen-Mammographie	Brustkrebs-Tod	360000 \$
Colestyramin (bei hohem Blutfell)	Infarkt-Tod	2000000 \$
Kostenlose medizin. Versorgung* für Männer)50 Jahre für Männer mit Risikofaktoren	Tod	ca. 600000 \$
	Tod	ca. 500000 \$
	Tod	ca. 130000 \$

* gilt für die USA

Obwohl unter Epidemiologen beliebt, sind meines Erachtens finanziellen Kosten ein schwieriges Argument: für die Krankenversicherungen und für einen Teil der Gesellschaft sind sie immer zu hoch, für einen anderen Teil der Gesellschaft, welcher vom wirtschaftswachstumskräftigen Gesundheitswesen und seiner Zuliefer-Industrie lebt, können sie nicht hoch genug sein.

Obiger Vergleich ist nicht erschöpfend und soll nur die Kuriosität aufzeigen, daß bei gängigen medizinischen Maßnahmen die Kosten pro verhindertem Tod in die Millionen gehen, und daß in finanzieller Hinsicht das Brustkrebs-Screening trotz allem dagegen noch billig erscheint. Nicht nur direkte finanzielle Kosten sind jedoch entscheidend, sondern auch Folgekosten und vor allem gesundheitliche Kosten.

Unerwünschte gesundheitliche und soziale Auswirkungen durch die Vorsorge

Was passiert mit den übrigen 1749 Frauen, die in 10 Jahren keinen Profit vom Screening erwarten können? Man kann mit einer Spekulation beginnen: Führt das 5000-fache Aufsuchen eines Mammographie-Zentrums nicht mit einer gewissen Wahrscheinlichkeit zu einem tödlichen Verkehrsunfall, welcher den Nutzen des gesamten Screenings einfach so schnell wieder wegblasen könnte? Tatsächlich war in beiden randomisierten Screenings-Studien die Gesamt-Todesrate in der Screening-Gruppe höher (Shapiro 1982, **Tabár** 1985b, Wright 1986):

Tabelle 8:

Brustkrebs- und Gesamt-Todesrate (pro 100000 Frauen-Jahre) in der schwedischen Studie und im HIP-Trial

Studie	Brustkrebs-Mortalität			Obrige Mortalität			Gesamt-Mortalität		
	SG	KG	Dijl.	SG	KG	Dijl.	SG	KG	Dijl.
Schweden +	15,9	21,6	(-5,7)	870,7	854,3	(+16,4)	886,6	875,9	(+10,7)
HIP New York++	45,8	60,2	(-14,4)	655,1	603,4	(+51,7)	700,9	633,6	(+37,3)

Studien kombinieren (Mantel-Haenszel); • =p=.11 (n.s.); OR = 1,03; •• =p=.025; OR = 1,04; + =von 7 Jahren (vgl. Tab. 5); •• =von 10 Jahren Beobachtung hochgerechnet

Die 30 %-ige Verringerung der Mammakarzinom-Todesfälle kann sich auf die Gesamt-Mortalität nicht signifikant auswirken, einfach **weil** nur ein geringer Teil ($\ll 10\%$) der Frauen insgesamt an Brustkrebs stirbt. Statistisch **zufällig** ist die Gesamt-Sterberate in den Screening-Gruppen sogar leicht höher ($p < .10$).

AuWilling ist aber die *statistisch signifikante* ($p < .05$) *Häufung aller nicht durch Brustkrebs verursachten Sterbe/älle in den Screeninggruppen* (eine 01' von 1.04 entspricht einem Anstieg von 4 %). Die »statistische Signifikanz« bedeutet, daß dieser Unterschied nicht durch Zufall bedingt ist. Durch das Screening könnten klinisch undramatische, subtile unerwünschte Wirkungen zum Zuge gekommen sein, die nun als Masseneffekte (vgl. auch Schmidt 1987a) in den 1749 Frauen mehr Mortalität erzeugen als der verhinderte Krebstod der einen Frau (in den Publikationen fehlen Angaben, um z.B. die Verkehrsunfall-Hypothese zu prüfen). Ein 4 %iger Unterschied kann aber eher als etwa ein 20 %-iger Unterschied vollständig durch einen bias (Verzerrung) verursacht sein und ich will obige Ergebnisse deshalb nicht überbewerten als Beweis, das Screening würde die **Nichtbrustkrebs-Sterblichkeit** signifikant erhöhen. Viele andere Einflüsse wie ein nicht kontrollierter Alterseffekt (mangelnde Alters-Standardisierung) oder unbekannte »Baseline«-Unterschiede müssen ebenfalls in Betracht gezogen werden. Außerdem sieht die Bilanz in einem weiteren Jahr Beobachtungszeit vielleicht schon wieder nicht mehr so dramatisch aus.

Dennoch wird mindestens die Plausibilität und Dringlichkeit der Hypothese belegt, das Screening könnte durch die massenhaften unerwünschten Effekte letztlich mehr Tod erzeugen als retten, und solche Hypothesen müßten deshalb in laufende Studien einbezogen werden. Leider wird durch die klinische Spezialisierung offenbar völlig vergessen, was die praktisch relevanten epidemiologischen Fragestellungen nach dem gesamthaften Leiden sind.

Tabelle 6 zeigte uns, wie problematisch ein Screening in Bezug auf massenhafte unerwünschte Gesundheits- und soziale Effekte ist: *Ein verhinderteter Krebstod wird erkauf mit:*

- Über 5000 Mammographien: Tausende Mal muß eine Frau die Prozedur der Mammographie auf sich nehmen, einige Tausend Stunden verlorener Freizeit oder Arbeitszeit, Nervosität, Verkehrsunfälle usw.?
- Über 200 zuerst verdächtigen Befunden: Hundertfache Verängstigung und um die 50 Nadel-Biopsien und Probe-Operationen ohne Konsequenzen auf einen kommenden oder nie kommenden Krebstod.
- Über 20 Frühentdeckungen: Nur in einem von 25 frühentdeckten Krebsen ist eine *zusätzliche* Heilung möglich. In allen anderen Fällen wird aber durch die Frühentdeckung ohne Nutzen die Krankheit schon früher »in Gang gestzt«, die Leidenszeit also verlängert.

Unerwünschte Vermehrung von Krebsfällen

Es läßt sich unterscheiden zwischen "Prävalenz-Screening« und »Inzidenz-Screening«. Am Anfang der Einführung eines Vorsorge-Programmes werden viele »schlummernde« kleine Krebse entdeckt, die in der Ziel-Gruppe bereits vorhanden sind und nun frühzeitig erfaßt werden (Prävalenz = pro-ef- Zeitpunkt vorhandene Fälle). Später werden dann durch das Screening zunehmend nur noch neuauftretende Krebse erfaßt (Inzidenz = pro Zeiteinheit neuauftretende Fälle). Wenn, wie aus *Tabelle 3* ersichtlich war, die Krebsrate durch das Screening zunimmt, so geht dies auf das Konto des Prävalenz-Screenings und bedeutet also nicht eine Krebsverursachung. Theoretisch sollte die Zahl neuauftretender Fälle mit der Zeit in Screening- und Kontroll-Gruppe identisch werden. Dem steht aber ein anderes theoretisches Argument entgegen. Obwohl das Screening fraglos keine Kreb&zell-Verursachung bedingt¹³, ist jedoch eine Krebsfall-Verursachung möglich.

Zur Epidemiologie und zur natural history des **Mammakarzinoms** gehört auch das Wissen, ob und wie häufig präklinische Krebsherde überhaupt zu einem Krankheitsausbruch führen. Dies ist bisher kaum untersucht worden. Vorläufige Daten sind nur aus einer Studie (Nielsen 1984) vorhanden, welche den sehr gutartigen Verlauf vieler (kleiner) Mammakarzinome aufzeigt. Bei 83 nacheinander durchgeführten Autopsien verstorbener Frauen (22-89 Jahre alt, mittleres Alter: 67) wiesen in der histologischen Brustuntersuchung 25 % (n=21) eine Form des Mammakarzinoms auf. Nur bei 29 % (n=6) dieser Frauen war das Karzinom jedoch vor dem Tode bekannt (und wahrscheinlich am Tode mit-

beteiligt). Eine histologisch »maligne« Kondition kann also sehr wohl (in 70 %!) klinisch gänzlich **gutartig** sein, ohne je Leiden zu verursachen, oder wegen anderer Todesursachen irrelevant bleiben.

Zu den im Screening fehl-diagnostizierten Karzinomen (17 % der Frühkarzinome, 6 % aller Karzinome [Working Group 1977]) kommen nun also noch eine ganze Reihe richtig diagnostizierter "Malignome«, die jedoch klinisch gutartig sind und somit im epidemiologischen Sinne falsch positiv. Die Sensibilität der modernen Technik hat also hier ganz beträchtliche Nachteile eingeführt.

Aus diesem Grund geht obige Liste noch weiter: Um die 10-20 Krebsfälle (pro verhütetem Todesfall) werden geschaffen, die ohne Screening nie klinisch in Erscheinung getreten wären!

Wie früh ist der kritische Interventions-Zeitpunkt?

Und es scheint, daß die hohe Sensibilität der modernen Mammographie die Heilbarkeit des Mammakarzinoms gar nicht verbessern konnte«Wir wollen dies zeigen, indem wir die Schwedische Studie dem HIP-Trial (Shapiro 1977, 1971, 1985) gegenüberstellen (Tabelle 9).

Tabelle 9:

Die Schwedische Studie und der HIP-Trial im Vergleich

		HIP-Trial New York	Swedische Studie
Beginn der Studie		1963	1977, 1978
Screening-Methode		Mammographie und Palpation	»Moderne« Mammographie
Teilnahme-Compliance		65 % (initial)	89 % (initial)
Altersdurchschnitt		51 (40-69)	57 (40-70)
Mammakarzinom-Inzidenz	SG:	2,0	2,3
	KG:	1,9	1,8
		(pro 1000 Frauenjahre)	
Prävalenz in erster Runde		2,7	5,0 (Kopparberg)
			(pro 1000 Teilnehmerinnen)
Metastasenfrei	SG:	57 %	73 %
Brustkrebs-Mortalität	SG:	2,6	1,6
	KG:	4,1	2,2
		(pro 10000 Frauenjahre)	
Senkung in %		37 %	31 %
Frauen unter 50		nicht signifikant	keine
Differenz		1,5	0,6
		(pro 10000 Frauenjahre)	
Unterschied bewahrt		Ja, über 14 Jahre	noch unbekannt

Obschon, wie aus der Tabelle ersichtlich, die Schwedische Studie im Früherkennen als solches beträchtlich mehr Erfolg hatte als die frühere HIP-Studie (Prävalenz in erster Screening-Runde, Anteil metastasenfreier Karzinome), ist der Erfolg beider Studien in Bezug auf die Mortalitätsreduktion etwa identisch. Die Hypothese, daß ein modernes Screening nun bei jüngeren Frauen unter 50, bei welchen nur sensible Methoden die schnellwachsenden Karzinome früh **entdecken** können, einen Effekt auf die Brustkrebs-Sterblichkeit zeigen würde (Wertheimer 1986), hatte sich nicht bewahrheitet.

Zudem zeigte die HIP-Studie auch das erstaunliche Resultat, daß der *mortalitätsverringernde Effekt* durch das Screening (ab dem fünften Jahr nach der Diagnose) *bei Lymphknoten-befallenen Patientinnen ausgeprägter* war als in Metastasen-freien Patientinnen (Shapiro 1985).

Diese verwirrenden Ergebnisse unterstreichen unsere Ignoranz über die natural history des Mammakarzinoms und lassen sich jedenfalls nicht einfach erklären. Es kann möglich sein, daß natürliche Unterschiede in der Biologie des Mammakarzinoms zwischen Frauen in Schweden und in den USA bestehen. Weiße Frauen haben Z.B. allgemein eine bessere Prognose als schwarze Frauen, auch wenn unterschiedliche sozio-ökonomische Bedingungen berücksichtigt werden (Bain 1986). Weiter sind auch Studien-Artefakte nie ganz auszuschließen, immerhin beträgt die drop-out-Rate im HIP-Trial nach 5 Jahren 25 %.

Eine mögliche und nicht sehr entfernte Hypothese ist aber auch, daß sich das Mammakarzinom tatsächlich in die eingangs postulierten zwei Gruppen unterteilt: Die einen gehören zum »lokalen Typ« und sind durch eine Früh-Intervention beeinflussbar, die anderen zum »systemischen Typ«. Scheinbar läßt sich der wirklich therapierbare lokale Typ aber auch noch im Stadium II, sogar mit regionärer Lymphknoten-Metastasierung, beeinflussen. Vor der Menopause treten möglicherweise vorwiegend nicht beeinflussbare systemische Karzinome auf. Dies könnte erklären, weshalb die dramatisch verbesserte Sensibilität der »Schwedischen« Technik gegenüber der alten Brustuntersuchung und unsensiblen Mammographie keine Vorteile (und offenbar nur Nachteile) gebracht hat. Die Faszination des herrschenden Mediziners am Diagnostizieren unterliegt vielen »fallacies« (Trugschlüssen), die sich auch in anderen Bereichen aufzeigen ließen (Schmidt 1987b).

Schlußfolgerungen

Berücksichtigen wir nun die gesammelte Evidenz, die Ergebnisse der zwei randomisierten Studien und die unveränderte Mammakarzinom-Sterblichkeit über die letzten Jahrzehnte, so läßt sich *theoretisch* feststellen, daß sich bei Früherkennung ein Teil der Mammakarzinome durch eine lokale Therapie heilen lassen und sich somit offenbar auch zuerst streng lokal entwickeln, bevor eine Metastasierung erfolgt. *Praktisch* kommt dem aber kaum Bedeutung zu. Die massenhaft wirksam werdenden unerwünschten Effekte bewirken eine Lebensqualitäts-Beeinträchtigung bei einer großen Zahl von Frauen und lassen es sogar möglich erscheinen, daß die Sterblichkeit insgesamt letztlich nicht verringert wird. Die Wahrscheinlichkeit, daß eine Frau weit häufiger zu einem unnötigen Krebsfall gemacht wird als daß sie von einer (zusätzlichen) Heilung profitieren kann, läßt die sensible mammographische Vorsorge-Untersuchung als unakzeptabel erscheinen. Die Idee, daß mit der Früherkennung als Gewinn schonendere Operationen verbunden sind, unterliegt somit zu einem guten Teil dem Unwissen der aufgezeigten Verfälschungseffekte klinischer Beobachtungen. In diesem Zusammenhang stellt sich auch ganz klar die Frage, ob nicht brusterhaltende Operationsverfahren die Regel werden sollten. Dies ist eine wichtige, schon zu lange verdrängte, Forschungsfrage. Das »weit-im-Gesunden-ausschneiden« muß revidiert werden und sollte von brustliebenden Menschen als wenig begründet in Frage gestellt werden.

Das frühe Diagnostizieren an sich ist kein Erfolg, der entscheidend ist. Die sensible moderne Mammographie weist gegenüber unsensibleren Methoden wie der Palpation offenbar keine eindeutigen Vorteile hinsichtlich der Verhinderung des Krebstodes auf, und es scheint, daß der kritische Interventionspunkt (critical point) entgegen gängiger Meinung sich in einem relativ späten Wachstums-Stadium befindet. Es scheint also, daß die Heilbarkeit bei **den** tatsächlich beeinflussbaren Karzinomen auch in einem relativ späten Stadium (II oder III) noch gegeben ist und sich durch eine möglichst frühe Intervention im Stadium des minimal cancers nicht verbessert. Diese Hypothese ist sicher bis zu einem gewissen Grade spekulativ, aber nicht mehr als die gängige Theorie. Gezielte Studien zur Abklärung dieser Frage sind notwendig, bevor definitive Aussagen möglich sind.

Praktisch wichtig ist aber auch die Feststellung, daß die hier dargestellten Parameter von absoluten Effekten (»attributable risks and benefits«) allein die Frage nach dem Nutzen beantworten können. So wie viele unlogisch konzipierte und mit Irrtums-Möglichkeiten behaftete

Studien noch immer verbreitet sind, so sind in der Interpretation von an sich gut konzipierten Studien logische Fehler im Gebrauch der epidemiologischen Parameter an der Tagesordnung.

Die simplistische Feststellung, mit dem Screening ließe sich die Brustkrebs-Todesrate senken und das Screening sei deshalb unbedingt nützlich, erweist sich als trügerisch und taugt ganz offensichtlich nicht als praktische Handlungsanweisung und ist als solche ein (epidemiologischer) Kunstfehler. Wir sind es den Frauen und Patientinnen schuldig, auf Nutzen *und Gefahren* des Screenings hinzuweisen. Werden die vielfältigen unerwünschten Effekte in ihrer Mehrzahl der Einzahl des verhüteten spezifischen Ereignisses (Brustkrebs-Tod) gegenübergestellt, so ist die Bilanz ernüchternd wenn nicht erschreckend.

Vom Screening ist deshalb abzuraten. Eine Ausnahme bleiben vielleicht die Selbstuntersuchung und die Mammographie bei Frauen mit hohem Brustkrebs-Risiko.

Anmerkungen

- Dieser Aufsatz beruht auf dem Arbeitspapier „Breast Cancer Screening - Critical Appraisal of the Evidence“, welches im Rahmen eines Weiterbildungskurses (1986) im Centre for Clinical Epidemiology, University of Newcastle Australia, zur Diskussion gestellt wurde. Obwohl der Aufsatz deshalb bis zu einem gewissen Grade „unabgerundet“ ist, eignet er sich aber gut, die Art und die Bedeutung klinischer-epidemiologischer Methoden in der Beurteilung vorhandener Forschungsergebnisse aufzuzeigen. Die vorläufige Schlussfolgerung aus dieser Analyse, daß (mammographische) Vorsorgeuntersuchungen eher mehr Leiden erzeugen als verhindern können, bleibt auch bei einer »abgerundeten« Arbeit bestehen (vgl. Schmidt J: Breast Cancer Screening - Is it Worthwhile?, wird voraussichtlich erscheinen im European Journal of Cancer)
- 2 Finanzielle Kosten-Überlegungen werden oft in den Vordergrund geschoben und es ist dabei dann folgerichtig, daß sich ein Screening für eine seltene Krankheit nicht lohnt. Dabei wird allerdings übersehen, daß „Kosten“ in Form von unerwünschten Wirkungen eines Vorsorge-Programmes ein viel größeres Problem darstellen können. Bei einer selteneren Krankheit ist der Untersuchungsaufwand nicht so sehr in finanzieller Hinsicht zu groß, sondern in Hinsicht auf die nötige Anzahl (nicht nebenwirkungsfreier) Untersuchungen und Folgeabklärungen, um einen Fall zu entdecken bzw. einen Krankheitsfall zu verhindern. So sind bei genauer Betrachtung auch in der heutigen „spärpolitischen“ Zeit Vorsorgeprogrammen eher Grenzen gesetzt durch deren zu hohe Kosten an unerwünschten Wirkungen als durch knapp gewordene Finanzen (vgl. auch Schmidt 1987a).
 - 3 Die „natural history“ einer Krankheit läßt sich etwa übersetzen mit: Der natürliche Krankheitsverlauf. Dieser Begriff der (klinischen) Epidemiologie enthält die Vorstellung, daß nur bei genauer Kenntnis der Natur und der möglichen Verläufe einer Krankheit überhaupt entschieden werden kann, ob »Heilungen« durch eine medizinische Maßnahme dieser oder einfach dem spontanen Verlauf zugeschrieben

ben werden können. Es ist dann beispielsweise wichtig zu wissen, ob histologisch »maligne« Gewebe tatsächlich regelmäßig als »bösartige« Krankheiten in Erscheinung treten.

- 4 Bei der randomisiert kontrollierten Studie handelt es sich um eine prospektive experimentelle bzw. Interventions-Studie. Die Studienpopulation wird dabei nach dem Los entweder der Kontrollgruppe oder der aktiven Studiengruppe, welche in unserem Fall zur Vorsorgeuntersuchung aufgefordert wird, zugeteilt. Aufgrund des Zufalles werden dabei prognostische Faktoren gleichmäßig auf beide Gruppen verteilt, und **Verfälschungen** durch Selbst-Auswahl werden verunmöglicht. Die Wirkung des Studienfaktors (hier Screening) und das Ausmaß dieser Wirkung lassen sich nach einer bestimmten Zeit direkt am Unterschied von **ee-Mortalitäts-Raten** oder anderen »Endpunkten« ablesen.
- 5 Bei der Fallkontroll-Studie handelt es sich um eine retrospektive Vergleichsstudie, bei welcher quasi vom »Endpunkt«, hier vom Mammakarzinom-Todesfall, ausgegangen wird. Für jeden Fall wird eine möglichst in allen Belangen (Alter etc.) identische Kontroll-Person bestimmt (»matching«), die jedoch frei ist vom entsprechenden Krankheitsereignis. Dann wird untersucht, ob bestimmte Faktoren, hier die Teilnahme am Screening, in Fall-Gruppe und Kontroll-Gruppe signifikant unterschiedlich verteilt sind. Dieser Studien-Typ ist relativ **wenig** aufwendig, kann jedoch kaum schlüssige Beweise einer tatsächlichen Assoziation von Krankheitsfaktoren und Krankheit liefern, da nicht alle Verfälschungsmöglichkeiten vermieden werden können.
- 6 Das Relative Risiko einer Intervention oder Exposition ist der Quotient der Inzidenz in der Studiengruppe dividiert durch die Inzidenz der Vergleichsgruppe. So haben die genannten Fallkontrollstudien ein RR von rund 0.5 ergeben (also 50 % Reduktion der Brustkrebs-Mortalität), während das **unverfälschte** RR aus randomisierten Studien rund 0.75 (25 % Reduktion) beträgt.
- 7 In Beobachtungsstudien werden in der Regel Häufigkeiten (Prävalenzen) von Krankheiten und Krankheitsstadien erfaßt (Querschnitts-Studie), wobei gleichzeitig vermutete Krankheitsfaktoren in die Erhebung einbezogen werden können. Das Auftreten von Krankheitsereignissen läßt sich auch im Längsschnitt beobachten (Kohorten-Studie), was aussagekräftiger ist (Zeit-Sequenz ist definiert). Beim Vergleich der Überlebenszeiten von screening-entdeckten Mammakarzinom-Fälle mit so augetretener Fälle wird jedoch weder die Querschnitts- noch die Kohorten-Studie richtig gebraucht. Jedenfalls operieren alle im Text erwähnten Verfälschungsmöglichkeiten in diesen »Studien«
- 8 Früherfaßte Karzinome bei Teilnehmerinnen (»verpaßte, falsch negative = Intervall-Karzinome).
- 9 Früherfaßte Karzinome in ganzer Zielgruppe (falsch negative = Intervall-Karzinome 1/2 Karzinome in Verweigerinnen). Da für das Individuum offensichtlich der Negativ Prädiktive Wert eher entscheidend ist, drückt die Sensibilität eher eine volksgesundheitlich relevante Größe aus. Von da her ist die »Erfassungskraft, des Screenings in der ganzen Zielgruppe (intenti on-to-screen) das aussagekräftigere Maß.
- 10 Der Positiv Prädiktive Wert (PPW) ist ein direktes Maß der Wahrscheinlichkeit, mit welcher Test-Positive eine Krankheit tatsächlich besitzen. Im praktischen Alltag wird häufig irrtümlicherweise angenommen, diese Wahrscheinlichkeit

würde durch die Test-Spezifität ausgedrückt. Die Spezifität ist jedoch definiert als der Anteil richtig als negativ klassifizierter Testlinge unter den tatsächlich krankheits-freien Testlingen. Dieses Maß ist für einen standardisierten Test konstant. Die Wahrscheinlichkeit, daß eine test-positive Person tatsächlich Krankheitsträger ist, wird jedoch genauso stark von der Prävalenz der Krankheit bestimmt wie von der Spezifität des verwendeten Testes. und entspricht nur bei einer Prävalenz von 1:2 genau der Spezifität (mit **einer** geringen Prävalenz wird auch der PPW niedrig, weil die Rate falsch positiver Resultate nun mit der größeren Wahrscheinlichkeit, von vornherein nicht krank zu sein, ansteigt).

- 11 Obwohl der relative Unterschied in der Metastasenhäufigkeit immerhin etwa 17 % beträgt, ist dieser statistisch gesehen hochgradig unsignifikant, also sehr wohl ein Zufallsprodukt. Eine Studie dieser Größe ist allerdings viel zu klein, um einen solchen Unterschied, falls er echt und nicht nur als Zufallsmöglichkeit bestehen sollte, zu erfassen (»power« = 20 %, Z-beta = +0.85). Ein statistisch nicht signifikantes Resultat bedeutet also nicht, daß eine Unterschiedslosigkeit bewiesen wäre (type I eTor). Um mit 90 %-iger Wahrscheinlichkeit den beobachteten Unterschied als statistisch signifikant (.05-level) zu erfassen, müßten insgesamt 400'000 Frauen in die Studie einbezogen werden. Dies zeigt, daß ein solcher Unterschied bei der letztlich doch geringen Inzidenz eines metastasierenden Karzinoms für eine einzelne teilnehmende Frau keine praktische Bedeutung hat. Wenn in einer derart großen Studie ein fast 20 %-iger Unterschied in der Metastasen-Häufigkeit statistisch nicht signifikant ist, so kann man **zwar** feststellen, daß daran zu 80 % eine zu kleine Probandenzahl schuld ist und ein Screening die Metastasierung vielleicht doch signifikant verhindert, wenn dies in fast einer halben Million Frauen untersucht worden wäre. **Ein** an der Ursachenfrage faszinierter Kliniker würde mit diesem Argument nochH nicht aufgeben wollen. Mir scheint, man müßte allerdings vielmehr folgern, daßs für die Perspektive eines Individuums ein Unterschied, welcher nur in einer halben Million Probanden als signifikant aufgezeigt werden kann, völlig irrelevant ist. Der beobachtete Unterschied besagt praktisch ausgedrückt tatsächlich, daß bei nur 1 von 1000 Frauen im Zeitraum von 10 Jahren (in 10000 Frauenjahren) durch die Vorsorge eine Metastasierung vielleicht verhütet wird. Auch wenn das »vielleicht« durch ein »signifikant« ersetzt werden könnte, bleibt diese Profit-Aussicht für ein Individuum doch bedeutungslos.
- 12 Nach dem eingangs entworfenen Modell des Mammakarzinom-Wachstums läßt sich aus der Verhütung der fortgeschrittenen Lokal-Pathologie nichts Definitives in Bezug auf die eventuelle Sterblichkeit ableiten. Obwohl sich diese Diskussion darüber deshalb eigentlich erübrigt, soll sie hier zur Uustration methodischer Fragen geführt werden. Das Zusammenfassen von »p-NX-Karzinomen« mit dem Stadium II ist mir in einer (persönlichen) Antwort des ersten Studien-Autors (17) auf mein (englisches) Manuskript als fachlicher Fehler angekreidet worden, weil ich nicht einmal die histologischen Stadien kennen würde; denn das »**p-NX**« sei im Grunde ein Stadium I-Karzinom. Das Bedeutsame an dieser professionellen Rüge ist das epidemiologisch zusammenhangslose Pochen auf herkömmlichem medizinischem Fachwissen (als autoritäres Gehabe). Epidemiologisch läßt sich das »fortgeschrittene Mammakarzinom« schlecht definieren, wie wir in den Ausführungen zur natural history gesehen haben, auch wenn saubere klinisch-histo-

logische Klassifizierungs-Schemen vorliegen. Das unklare .p-NX.-Stadium kommt jedenfalls in der Screening-Gruppe viel häufiger vor und es ist deshalb geboten, auch einen Vergleich anzustellen für alle fortgeschrittenen und potentiell fortgeschrittenen Karzinome, also .fortgeschritten. definiert als .nicht eindeutig Stadium I«. Spätere, bisher unpublizierte Daten, die mir die Autoren zur Verfügung stellten, zeigen im etwas längeren Verlauf auch für die Metastasierungsrate statistisch signifikante Unterschiede. Aber auch hier ist schon wieder Vorsicht am Platz. Nach statistischer Lehre müssen bei wiederholten Vergleichen Signifikanz-Kriterien stringenter definiert werden. Schaut man nämlich 20mal, ob ein Unterschied auf dem .05-level signifikant ist, dann ist rein durch Zufall einmal (20 % .05) ein (vermeintlich) signifikantes Resultat zu finden. Die fragliche Signifikanz mahnt deshalb vor Überbewertungen dieses Ergebnisses.

- 13 Die Strahlenbelastung der modernen Mammographie ist gering und es wird geschätzt, daß lediglich 1 % aller Mammakarzinome durch ein Screening verursacht würden (Fitzgerald 1983, Zuur 1985).

Literaturverzeichnis

- Bailar III JC and Smith EM: „Progress Against Cancer?“, N Engl J Med 314:1226, 1986
- Bain RP et al: »Racial Differences in Survival of Women with Breast Cancer“, J Chron Dis 39:631, 1986
- Behars OH et al: »Report of Working Group to Review BCDDP“, J Natl Cancer Inst 62:679, 1979
- Coilette HJA et al: »Evaluation of Screening for Breast Cancer in a Non-Randomized Study (the DOM project) by Means of a Case-Control Study“, Lancet 1:1224, 1984
- Editorial »Early Diagnosis and Survival in Breast Cancer«, Lancet 11:785, 1981
- Fitzgerald M: »Radiation Hazard of Breast Cancer Screening“ >BJ J Radial 56:283, 1983
- Haagensen CD et al: . Breast Carcinoma — Risk and Detection“, WB Saunders, Philadelphia 1981
- Henderson C and Canellos GP: »Cancer of the Breast - The Past Decade“, N Engl J Med 302:17 + 78, 1980
- Hendricks JHCL and Verbeek ALM: »Population Screening for Breast Cancer by Mammography in The Netherlands - Expectations, Early Results, Negative Effects and Conditions for Large-Scale Screening.« >Diagn Imag din Med 54:186, 1985
- Himmelstein DU and Woolhandler S: »Free Care, Cholestyramine, and Health Policy“, N Engl J Med 311:1511, 1984
- Miller AB et al: »The National Study of Breast Cancer Screening. Protocoll for a Canadian Randomized Controlled Trial of Screening for Breast Cancer in Women., Clin Invest Med 4:227> 1981
- Nielsen M et al: . Precancerous and Cancerous Breast Lesions During Lifetime and al Autopsy“, Cancer 54:612, 1984
- Redding WH et al: .Delection of Micrometastases in Patients with Primary Breast Cancer“, Lancet 11:1271, 1983

- Roberts MM et al: »The Edinburgh Randomized Trial of Screening for Breast Cancer: Description of Method«, *Br J Cancer* 50:[, 1984
- Rodes ND et al: »The Impact of Breast Cancer Screening on Survival. A 5- to [0-Year Follow-Up Study., *Cancer* 57:58[, 1986
- Sackett DL et al: »Clinical Epidemiology - A Basic Science for Clinical Medicine., p.139ff., Little, Brown and Company, Boston/Toronto 1985
- Schmidt J: .Risiko und Nutzen einer Behandlung: Das Beispiel C10fibrat«, *Schweiz Med Wschr* 117:773, 1987
- Schmidt J: .Screening for Aids, Can it Ever Make Sense?«, *Manuskript* 1987
- Shapiro. S et al: »Periodic Breast Cancer Screening in Reducing Mortality From Breast Cancer«, *J Am Med Ass* 2[5:1777, [971
- Shapiro S: .Evidence on Screening for Breast Cancer from a Randomized Trial«, *Cancer* 39(suppl):2772, [977
- Shapiro S et al: .Ten to Fourteen Year Effect of Screening on Breast Cancer Mortality«, *J Natl Cancer Inst* 69:349, 1982
- Shapiro S et al: »Selection, Follow-Up, and Analysis in the HIP Study: A Randomized Trial With Breast Cancer Screening«, *Natl Cancer Inst Monogr* 67:65, 1985
- Tabár L et al: .Significant Reduction in Advanced Breast Cancer - Results of the First Seven Years of Mammography Screening in Kopparberg, Sweden«, *Diagn Imag clin Med* 54:158, 1985
- Tabár L et al: .Reduction in Mortality from Breast Cancer After Mass Screening with Mammography«, *Lancet* 1:829, 1985
- Verbeek ALM et al: .Reduction of Breast Cancer Mortality Through Mass Screening With Modern Mammography. First Results of the Nijmegen Project 1975-81«, *Lancet* 1:1222, 1984
- Wertheimer MD et al: .Increasing the Effort Toward Breast Cancer Detection«, *J Am Med Ass* 255:[311, 1986
- Working Group to Review the NCIACS Breast Cancer Demonstration. Detection Projects, Report. National Cancer Institute, National Institute of Health, Bethesda Maryland 1977
- Wright CJ: .Breast Cancer Screening: A Different Look at the Evidence., *Surgery* 100:594, [986
- Zuur C and Broerse I: .Risk and Cost-Benefit Analysis of Breast Cancer Screening Programs Derived from Absorbed Dose Measurements in the Netherlands«, *Diagn Imag clin Med* 54:211, [985