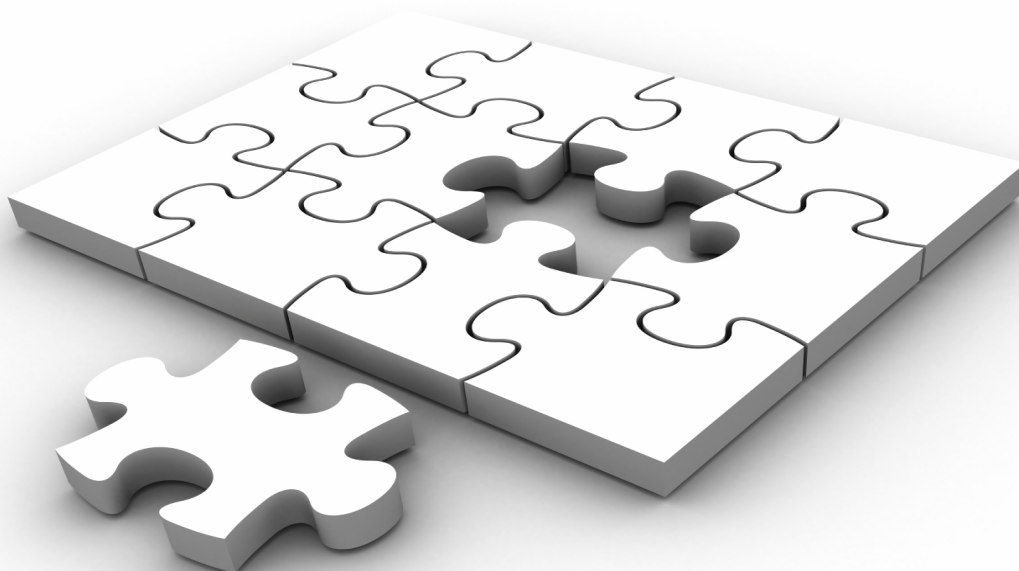


Evidenzbericht 2007 zur S-3-Leitlinie Brustkrebsfrüherkennung in Deutschland

**Systematische Suche nach Informationen zum
medizinisch-wissenschaftlichen Kenntnisstand
und Bewertung der Evidenz zur Aktualisierung
und Überarbeitung**

Version 1.00 / August 2007



© Ärztliches Zentrum für Qualität in der der Medizin (ÄZQ)

Impressum

Herausgeber

Ärztliches Zentrum für Qualität in der Medizin (ÄZQ)

Gemeinsame Einrichtung von Bundesärztekammer und
Kassenärztlicher Bundesvereinigung

Titel

Evidenzbericht 2007 zur S-3-Leitlinie Brustkrebsfrüherkennung in Deutschland. Systematische Suche nach Informationen zum medizinisch-wissenschaftlichen Kenntnisstand und Bewertung der Evidenz zur Aktualisierung und Überarbeitung.

Auftraggeber

Konzertierte Aktion zur Brustkrebsfrüherkennung, vertreten durch PD Dr. Ute-Susan Albert, Universitätsklinikum Marburg, für die Arbeitsgruppen zur Aktualisierung der S-3-Leitlinie Brustkrebsfrüherkennung in Deutschland

Autoren / Bearbeitung

Dr. Monika Nothacker, MPH, Ärztliches Zentrum für Qualität in der Medizin, Berlin (Koordination)

Monika Lelgemann MSc, HTA Zentrum in der Universität Bremen (Teil A)

Dr. Klaus Giersiepen MPH, Bremer Institut für Präventionsforschung und Sozialmedizin (Teil A)

Dr. Susanne Weinbrenner MPH, Ärztliches Zentrum für Qualität in der Medizin, Berlin

Wir danken den beteiligten Experten der konzertierten Aktion zur Brustkrebsfrüherkennung für ihre inhaltlichen Beiträge: Frau Professor I. Schreer (Mammographie-Screening und Doppelbefundung), Frau PD Dr. A. Lebeau (Her-2/neu-Bestimmung), Herrn Dr. Duda (Mammasonographie) sowie Herrn Prof. R. Schulz-Wendtland und Herrn PD Dr. M. Müller-Schimpfle (Stellenwert der Vakuumbiopsie/ Tumorzellverschleppung) sowie Frau PD U.-S. Albert für die Koordination.

Für die methodische Begleitung des Gesamtberichts danken wir Frau Monika Lelgemann, Msc.

Der Bericht wurde im August 2007 fertiggestellt. Die Aussagen wurden auf Grundlage der zum jeweiligen Zeitpunkt zur Verfügung stehenden und in die Auswertung einbezogenen Studien getroffen.

Redaktion und Gestaltung

Dipl.-Soz.Päd. Marga Cox, Ärztliches Zentrum für Qualität in der Medizin, Berlin

Anschrift des Herausgebers

Ärztliches Zentrum für Qualität in der Medizin (ÄZQ)

Wegelystraße 3 / Herbert-Lewin-Platz

10623 Berlin

Tel 030 / 4005 – 2500

E-Mail: mail@azq.de

www.azq.de

© August 2007 – Ärztliches Zentrum für Qualität in der Medizin

Inhalt

I. HINTERGRUND / ZUSAMMENFASSUNG DER ERGEBNISSE

1. Auftrag.....	5
2. Themenbereiche / Fragestellungen	5
3. Vorgehensweise	7
4. Ergebnisse	8

II. THEMENBEREICHE

A. Mammographie-Screening.....	12
B. Doppelbefundung von Screening-Mammographien.....	36
C. Stellenwert der Mammasonographie im Rahmen der Brustkrebsfrüherkennung.....	80
D. Indikationsstellung für die stereotaktische Vakuumbiopsie im Vergleich zur stereotaktischen Stanzbiopsie.....	148
E. Nutzen der Exzision des Stichkanals nach minimal-invasiver Biopsie der Brust zur Vermeidung von Tumorzellverschleppung.....	172
F. Bewertung der Her-2-neu-Bestimmungsmethoden.....	192

III. ANHANG

Anhang 1. Oxford Levels of Evidence.....	274
--	-----

Abkürzungsverzeichnis

ACP	American College of Physicians
ACR	American College of Radiology
ADH	Atypische duktale Hyperplasie
AK	Anitkörper
AS	Axillasonographie
ÄZQ	Ärztliches Zentrum für Qualität in der Medizin
Bef.	Befunder
BI-RADS TM	Breast Imaging Reporting and Data System
CAD	Computer-Assistierte-Diagnose
CA	Carcinoma
CBE	Clinical Breast Examination (professionelle Palpation der Brust)
CISH	Chromogene-in-situ-Hybridisierung
CNB	Core Needle Biopsy (Stanzbiopsie)
CNBSS	Canadian National Breast Screening Study
CRD	Centre for Reviews and Dissemination
CTA	Clinical Trial Assay
DCIS	Ductal Carcinoma in-Situ (Duktales In Situ Karzinom)
DEGUM	Deutsche Gesellschaft für Ultraschall in der Medizin
EUSOMA	European Society of Mastology (Europäische Gesellschaft für Brustkunde)
FISH	Fluoreszenz-in-situ-Hybridisierung
FNA	Feinnadelaspiration
FNAC	Fine Needle Aspiration Cytology
FNB	Feinnadelbiopsie
FSE	Final Surgical Excision (Definitives Operationspräparat)
Her2	Human Epidermal Growth Factor Receptor 2
IC	Invasive Carcinoma (Invasives Karzinom)
IG	Image Guidance
IHC	Immunhistochemie
KI	Konfidenzintervall
L/T- Ratio	Length to Thickness - Ratio
LK	Lymphknoten
MA	Meta-Analysen
Mhz	Megahertz
MM	Mammographie
MRT	Magnetresonanztomographie
MS	Mammosonographie
NNH	Number Needed to Harm
NNS	Number Needed to Screen
NPV	Negativer prädiktiver Value (Wert)
NTS	Needle Track Seeding
p	p-Wert (Irrtumswahrscheinlichkeit)
PPV	Positiver prädiktiver Value (Wert)

RADS	R eporting a nd D ata S ystem
RCT	R andomized C ontrolled T rial
ROC	R eceiver- O perating- C haracteristics
RR	R elatives R isiko
RTA	R öntgenassistent
SNB	S entinel- N ode- B iopsy (Biopsie des Wächterlymphknotens)
Sono	S onographie (Ultraschall)
SR	S ystematischer R eview (Systematische Übersichtsarbeit)
T-Stadien	T umor-Stadien
V.a.	V erdacht a uf
VAB	V akuumbiopsie
WL	W ire L ocalization
ACP	A merican C ollege of P hysicians
ACR	A merican C ollege of R adiology

1. Auftrag

Anlässlich der Aktualisierung der erstmals 2003 erschienenen S-3-Leitlinie Brustkrebsfrüherkennung der Konzentrierten Aktion zur Brustkrebsfrüherkennung wurden im Rahmen eines formalisierten Konsensverfahrens Themenbereiche bzw. Schlüsselfragen identifiziert, für die eine systematische Aufbereitung der Literatur erfolgen sollte. Der Auftrag zur Bearbeitung der aufgeführten Fragen erging im August 2006 an das Ärztliche Zentrum für Qualität in der Medizin (ÄZQ) in Berlin, die Themenbereiche wurden bis zum Januar 2007 durch die Expertenarbeitsgruppen konkretisiert.

2. Themenbereiche / Fragestellungen

A. Mammographie-Screening

1. Besteht ein nachgewiesener Nutzen eines Mammographie-Screenings für Frauen im Alter von 40-49 Jahren?
2. Besteht ein nachgewiesener Nutzen eines Mammographie-Screenings für Frauen älter als 70 Jahre?

B. Doppelbefundung von Screening-Mammographien

1. Welchen diagnostischen Nutzen bzw. Schaden hat die Doppelbefundung von Screening-Mammographien unter Berücksichtigung der verschiedenen Entscheidungsverfahren bei nicht übereinstimmenden Befunden?
2. Welchen diagnostischen Nutzen bzw. Schaden hat der Einsatz von Computer-Assistierten-Diagnose-Systemen im Rahmen des Mammographie-Screenings?

C. Stellenwert der Mammasonographie im Rahmen der Brustkrebsfrüherkennung

1. Kann eine zusätzliche Mammasonographie beim Brustkrebs-Screening die Karzinomentdeckung verbessern? Welcher potentielle Schaden entsteht für die untersuchten, asymptomatischen Frauen?
2. Welche diagnostische Sicherheit weist die Sonographie im Hinblick auf die Beurteilung axillärer Lymphknoten bei negativem klinischem Befund auf?
3. Welchen Nutzen hat die Mammasonographie für die Karzinomdiagnostik bei auffälligem Brustbefund?
4. Welchen Stellenwert hat die Mammasonographie bei der Abklärung auffälliger Brustbefunde bei Frauen unter 50 Jahren?

D. Indikationsstellung für die stereotaktische Vakuumbiopsie im Vergleich zur stereotaktischen Stanzbiopsie

1. Für welche mammographischen Befunde ist die stereotaktische Vakuumbiopsie der stereotaktischen Stanzbiopsie hinsichtlich der diagnostischen Sicherheit überlegen?
2. Welche Unterschiede hinsichtlich Art und Umfang von Komplikationen treten bei den beiden Verfahren auf?

E. Nutzen der Exzision des Stichkanals nach minimal-invasiver Biopsie der Brust zur Vermeidung von Tumorzellverschleppung

1. Ist die Exzision des Stichkanals nach minimal-invasiver Biopsie aufgrund einer möglichen metastastischen Tumorabsiedelung erforderlich?

F. Bewertung der Her-2-neu-Bestimmungsmethoden

1. Sprechen Patientinnen mit Her2-Rezeptor-Überexpression auf der Zelloberfläche in gleichem Maße auf eine Therapie mit Trastuzumab an wie Patientinnen mit einer vermehrten Expression des entsprechenden Gen-Abschnitts auf Chromosom 17?
2. Inwieweit sind die Methoden Immunhistochemie (IHC) und Fluoreszenz in situ Hybridisierung (FISH) reliabel ?
3. Sind die Methoden IHC und FISH im Hinblick auf ein negatives (0,1+) oder positives (3+) Ergebnis konkordant?
4. Inwieweit ist die Methode Chromogene in situ Hybridisierung reliabel und konkordant mit der FISH-Bestimmung?
5. Entspricht die diagnostische Wertigkeit der Her2- Bestimmung durch IHC / FISH / CISH aus Biopsiematerial der einer Bestimmung aus dem definitiven Operationspräparat?

3. Vorgehensweise

Nach Einarbeitung in die Thematik erfolgte die systematische Bearbeitung der aufgeführten Fragen entsprechend dem unten stehenden Schema. Durch die Fragestellungen eventuell bedingte Modifikationen des Vorgehens sowie detaillierte Angaben sind den einzelnen Kapiteln zu entnehmen.

Abbildung 1
Erstellung des Evidenzberichts - Vorgehen



4. Ergebnisse

A. Mammographie-Screening

- 1) *Besteht ein nachgewiesener Nutzen eines Mammographie-Screenings für Frauen im Alter von 40-49 Jahren?*

Für Frauen der Altersgruppe 40-49 Jahre sprechen die vorliegenden Daten für einen geringen relativen Nutzen (15%) bezüglich des Endpunktes „Brustkrebssterblichkeit“.

Der absolute Nutzen bezüglich des Endpunktes Brustkrebssterblichkeit liegt bei etwa einer von ca. 2.500 Frauen, die dazu über einen Zeitraum von 10 Jahren regelmäßig zum Screening eingeladen werden müssen. Dem stehen falsch-positive Befunde kumulativ über 10 Jahre bei 20-49% aller Eingeladenen gegenüber, dazu eine noch ungewisse Anzahl von Überdiagnosen sowie möglicherweise eine dem Nutzen vergleichbare Anzahl von zusätzlich induzierten Mammakarzinomen.

Hinsichtlich der möglichen negativen Effekte besteht aufgrund spärlicher und schlecht übertragbarer Datenlage eine große Unsicherheit der Aussagen.

Dementsprechend gilt diese Unsicherheit auch für die Beurteilung des so genannten Netto-Nutzens.

Die Entscheidung für oder gegen ein Mammographie-Screening in dieser Altersgruppe wird aufgrund der dargestellten Studienlage stark von wertenden Abwägungen bestimmt werden

- 2) *Besteht ein nachgewiesener Nutzen eines Mammographie-Screenings für Frauen älter als 70 Jahre?*

Für Frauen der Altersgruppe 70-74 Jahre sprechen die vorliegenden Daten eher für einen Schaden als für einen Nutzen bezüglich des Endpunktes „Brustkrebssterblichkeit“.

Die Datenlage zu dieser Fragestellung ist allerdings sowohl hinsichtlich des potentiellen Nutzens als auch des potentiellen Schadens sehr spärlich, so dass alle Entscheidungen mit einer großen Unsicherheit behaftet sind. Den vorliegenden Daten kann kein hinreichender Beleg für einen Nutzen entnommen werden, so dass ein Mammographie-Screening in dieser Altersgruppe nicht gerechtfertigt scheint.

B. Doppelbefundung von Screening-Mammographien

- 1) *Welchen diagnostischen Nutzen bzw. Schaden hat die Doppelbefundung von Screening-Mammographien unter Berücksichtigung der verschiedenen Entscheidungsverfahren bei nicht übereinstimmenden Befunden?*

Die vorliegenden Daten zeigen, dass sich die Sensitivität des Mammographie-Screenings durch die Doppelbefundung jeglichen Entscheidungsverfahrens um 2,9%-13,7% (im Median um 7,8%) erhöht.

Die Spezifität durch Doppelbefundung bleibt gleich oder erhöht sich um bis zu 2,8% bei Einigung beider Befunder bei diskrepanten Befunden oder Entscheidung durch einen Dritten (Konsensus- oder Arbitration-Verfahren). Bei Berücksichtigung aller auch nur durch einen Befunder als suspekt befundenen Mammographien (unilaterales Verfahren) verschlechtert sich die Spezifität. Im Hinblick auf einen Schaden durch Doppelbefundung ist festzustellen, dass die Rate unnötig wiedereinbestellter Frauen bei Konsensus- und Arbitrationverfahren bei höherer Sensitivität des Screenings eher gesenkt wird, während sie bei unilateralem Verfahren steigt.

Keine der vorliegenden Studien verfolgte den Endpunkt Brustkrebsmortalität, so dass Nutzen bzw. Schaden nicht abschließend bewertet werden können.

- 2) *Welchen diagnostischen Nutzen bzw. Schaden hat der Einsatz von Computer-Assistierten-Diagnose-Systemen (CAD) im Rahmen des Mammographie-Screenings?*

Für den Einsatz von CAD-Systemen als Ersatz für einen Zweitbefunder lagen nur wenige Daten mit direktem Vergleich zur Doppelbefundung vor. Diese zeigen im Hinblick auf die Rate entdeckter Karzinome widersprüchliche Ergebnisse. Aufgrund der Studienlage kann derzeit nicht geschlossen werden, dass CAD-Systeme den menschlichen Zweitbefunder ersetzen können. Bei dem Einsatz von CAD-Systemen erhöht sich grundsätzlich die Rate wiedereinbestellter Frauen. Weitere Bewertungen zu Nutzen und Schaden sind aufgrund mangelnder Daten derzeit nicht möglich.

C. Stellenwert der Mammasonographie im Rahmen der Brustkrebsfrüherkennung

- 1) *Kann eine zusätzliche Mammasonographie beim Brustkrebs-Screening die Karzinom-entdeckung verbessern? Welcher potentielle Schaden entsteht für die untersuchten, asymptomatischen Frauen?*

Die vorliegenden Studien geben einen Hinweis auf einen zusätzlichen Nutzen der Mammasonographie nach negativer Mammographie im Rahmen eines Brustkrebs-Screenings bei Frauen mit dichter Brust (vorwiegend BI-RADS Klassen 3 und 4). Es wurden überwiegend invasive Karzinome von einer medianen Größe bis 1cm diagnostiziert, im Mittel 3,6 pro tausend Frauen dieser Gruppe. Da die untersuchten Altersgruppen eine sehr große Altersspanne aufwiesen und durchgehend auch jüngere Frauen miteinbezogen waren, als es für das Screening (derzeit 50-69 Jahre) gilt, sind prospektive Validierungsstudien im Rahmen eines realen Screening-Programms erforderlich. Im Hinblick auf den Schaden durch die Intervention Mammasonographie führte eine hohe Rate an falsch positiven sonographischen Befunden zu einer mehr als doppelt so hohen Rate an Biopsien im Vergleich zu Screening Mammographien. Eine abschließende Beurteilung des Nutzens wie auch des Schadens ist aufgrund mangelnder Daten nicht möglich.

- 2) *Welche diagnostische Sicherheit weist die Sonographie im Hinblick auf die Beurteilung axillärer Lymphknoten bei negativem klinischem Befund auf?*

Die Sensitivität der axillären Sonographie weist eine hohe Variabilität auf (26,4-94,5%), ebenso wie der negative prädiktive Wert (0,64-0,93). Die Vergleichbarkeit der Daten wird erschwert durch die Anwendung unterschiedlicher sonographischer Malignitätskriterien.

In der Gesamtbeurteilung der Studienlage kann die unauffällige Axillasonographie bei negativem klinischem Befund karzinomatös befallene Lymphknoten derzeit nicht mit so großer Sicherheit ausschließen, dass auf eine weitere operative Abklärung verzichtet werden kann.

- 3) *Welchen Nutzen hat die Mammasonographie für die Karzinomdiagnostik bei auffälligem Brustbefund?*

Die vorliegenden Studien zeigten bei auffälligen Brustbefunden in Bezug auf die Karzinomentdeckung eine verbesserte Sensitivität von 6%-30,9% bei zusätzlichem Einsatz der Mammasonographie verglichen mit der alleinigen Mammographie. Insgesamt sind die Daten als zu heterogen zu bewerten, um den Nutzen der Mammasonographie quantitativ valide zu bewerten. Zusammenfassend hat die Mammasonographie bei palpablen Befunden eine zur Mammographie ergänzende Rolle, diese ist jedoch nicht spezifisch für die Brustkrebsfrüherkennung.

- 4) *Welchen Stellenwert hat die Mammasonographie bei der Abklärung auffälliger Brustbefunde bei Frauen unter 50 Jahren?*

Für die Frauen unter 50 Jahren zeigten die vorliegenden Daten im Vergleich von Mammasonographie und Mammographie durchgehend eine bessere Sensitivität für die Mammasonographie. Aufgrund unzureichender Daten kann keine Altersgrenze für die Empfehlung des primären Einsatzes der Mammasonographie angegeben werden.

D. Indikationsstellung für die stereotaktische Vakuumbiopsie im Vergleich zur stereotaktischen Stanzbiopsie

- 1) *Für welche mammographischen Befunde ist die stereotaktische Vakuumbiopsie der stereotaktischen Stanzbiopsie hinsichtlich der diagnostischen Sicherheit überlegen?*

Die Frage nach der Indikationsstellung für die stereotaktische Vakuumbiopsie kann nur aus Einzelstudien mit insgesamt eher geringerer statistischer Aussagekraft beantwortet werden. Im Vergleich zur stereotaktischen Stanzbiopsie ergibt sich für den mammographischen Mikrokalkbefund in der Zusammenschau ein deutlicher Vorteil für die Vakuumbiopsie, nicht aber bei Herdbefunden.

Bei Mikrokalk zeigt sich sowohl für die Rate falsch-negativer Befunde wie auch für die Rate an Unterdiagnostik eine Differenz von jeweils über 10% zugunsten der stereotaktischen Vakuumbiopsie.

Die vorliegenden Meta-Analysen beantworten die Frage nach der Überlegenheit der stereotaktische Vakuumbiopsie nur durch indirekte Vergleiche, ohne Analyse der mammographisch vorliegenden Morphologie der Befunde. Die Rate falsch negativer Befunde ist bei beiden Verfahren gering (max. 3,1%). Die stereotaktische Vakuumbiopsie bietet hier einen Vorteil von 0,8%-1,6%. Im Hinblick auf die Unterdiagnostik von Malignomen (in-situ-Karzinomen und invasives Karzinom) aus Befunden atypischer duktaler Hyperplasie zeigen die Meta-Analysen eine deutliche Überlegenheit der stereotaktischen Vakuumbiopsie, für die Unterschätzung von invasiven Karzinomen allein zeigen sich widersprüchliche Ergebnisse.

Aufgrund der Datenlage können die Ergebnisse nicht als definitiv valide bezeichnet werden.

- 2) *Welche Unterschiede hinsichtlich Art und Umfang von Komplikationen treten bei den beiden Verfahren auf?*

Die insgesamt angegebenen Komplikationen lagen für die stereotaktische Stanzbiopsie bei 0,6% und für die stereotaktische Vakuumbiopsie bei 1,3%. Aufgrund lückenhafter Angaben zu Komplikationen in den Studien ist die Validität der Ergebnisse eingeschränkt.

E. Nutzen der Exzision des Stichkanals nach minimal-invasiver Biopsie der Brust zur Vermeidung von Tumorzellverschleppung

- 1) *Ist die Exzision des Stichkanals nach minimal-invasiver Biopsie aufgrund einer möglichen metastastischen Tumorabsiedelung erforderlich?*

Zusammenfassend lässt sich festhalten, dass es bisher keine Evidenz für die Empfehlung zur Exzision des Stichkanals gibt. Allerdings kann auch nicht ausgeschlossen werden, dass es aufgrund einer fehlenden Exzision des Stichkanals zu lokalen Rezidiven kommt, wie sie in Einzelfällen beschrieben sind.

F. Bewertung der Her-2-neu-Bestimmungsmethoden

- 1) *Sprechen Patientinnen mit Her2-Rezeptor-Überexpression auf der Zelloberfläche in gleichem Maße auf eine Therapie mit Trastuzumab an wie Patientinnen mit einer vermehrten Expression des entsprechenden Gen-Abschnitts auf Chromosom 17?*

Insgesamt kann aus der vorliegenden Datenlage nur eine sehr eingeschränkt valide Aussage zum Therapieansprechen getroffen werden. Die derzeitige Vorgehensweise, sowohl IHC3+ als auch FISH positive Patientinnen für eine Therapie mit Trastuzumab zuzulassen, wird durch die Studien in der Tendenz bestätigt. Um ein sicheres Testprocedere festzulegen, sind weitere Studien mit höherer statistischer Aussagekraft im Hinblick auf das Ansprechen von Trastuzumab erforderlich, insbesondere mit den jetzt gebräuchlichen Antikörpern.

- 2) *Inwieweit sind die Methoden Immunhistochemie (IHC) und Fluoreszenz in-situ-Hybridisierung (FISH) reliabel ?*

Die reliablere Methode sowohl hinsichtlich der Übereinstimmung zwischen mehreren Befundern als auch Laboren ist der FISH-Test. Die immunhistochemische Bestimmung zeigt eine große Variabilität, auch für die Reliabilität bei Verwendung gleicher Antikörper.

- 3) *Sind die Methoden IHC und FISH im Hinblick auf ein negatives (0,1+) oder positives (3+) Ergebnis konkordant?*

Die meisten Studien verwendeten den Herceptest. Für diese Studien ergab die berechnete gepoolte Konkordanz von IHC 3+ zu FISH einen Wert von 88,09% (1724/1957).

Die gepoolte Konkordanz von IHC 0/1+ zu FISH betrug 96,1% (6105/6354).

In der Gesamtwertung zeigten die Kategorien IHC0/1+ eine sehr hohe Konkordanz und die Kategorie IHC3+ eine vertretbare Konkordanz.

- 4) *Inwieweit ist die Methode Chromogene in-situ-Hybridisierung reliabel und konkordant mit der FISH-Bestimmung?*

Die angegebenen Kappa-Werte sprechen einheitlich für eine exzellente Konkordanz zwischen den beiden Verfahren. Für die Diskriminierung der Her-2/neu Genamplifikation in positiv und negativ können das CISH-Verfahren und das FISH-Verfahren demzufolge als gesichert konkordant gelten.

Im Hinblick auf die Reliabilität der CISH liegen bisher nur wenige Daten vor. Diese zeigen eine der FISH vergleichbare Reliabilität.

- 5) *Entspricht die diagnostische Wertigkeit der Her2-Bestimmung durch IHC / FISH / CISH aus Biopsiematerial der einer Bestimmung aus dem definitiven Operationspräparat?*

Die vorliegenden Daten zeigten insgesamt keine konsistente Übereinstimmung der immunhistochemischen Bestimmungsergebnisse für Her-2/neu aus Biopsiematerial im Vergleich zum definitiven Operationspräparat.

Für die Methode FISH liegen nur Daten aus einer Studie vor. Aufgrund der Datenlage kann die immunhistochemische Bestimmung aus Biopsiematerial aufgrund der nicht sicheren Übereinstimmung mit dem definitiven Operationspräparat derzeit nicht empfohlen werden.

Weitere prospektive Studien sind für eine valide Beurteilung der Übereinstimmung von Biopsieergebnis und Ergebnis am definitiven Operationspräparat für beide der Bestimmungsmethoden erforderlich.

A

Mammographie-Screening bei Frauen 40-49 Jahre und >70 Jahre Nutzen/Schaden-Bilanz

B

Doppelbefundung von Screening-Mammographien unter Berücksichtigung des Entscheidungsverfahrens bei diskrepanten Befunden und des Einsatzes von Computer-Assistierten-Diagnose (CAD)-Systemen

C

Stellenwert der Mammasonographie im Rahmen der Brustkrebsfrüherkennung

D

Indikationsstellung für die stereotaktische Vakuumbiopsie im Vergleich zur stereotaktischen Stanzbiopsie

E

Nutzen der Exzision des Stichkanals nach minimal-invasiver Biopsie der Brust zur Vermeidung von Tumorzellverschleppung

F

Her-2/neu-Bestimmung:
Vergleich der Methoden Immunhistochemie, Fluoreszenz-in-situ-Hybridisierung und Chromogene-in-situ-Hybridisierung

Inhalt

1. Hintergrund.....	15
1.1 Screening.....	15
1.2 Nutzen.....	16
2. Fragestellung.....	16
3. Methodik.....	16
3.1 Kriterien für den Einschluss von Studien.....	16
3.1.1 Population.....	16
3.1.2 Intervention und Vergleich.....	16
3.1.3 Studientypen.....	17
3.2 Literaturrecherche.....	17
3.3 Endpunkte.....	18
3.3.1 Nutzen/Schaden –Bilanz.....	19
4. Ergebnisse Frauen 40-49 Jahre.....	20
4.1 Evidenzgrundlage.....	20
4.1.1 Ergebnis der Literaturrecherche.....	21
4.2 Nutzen.....	21
4.2.1 Brustkrebsmortalität.....	21
4.2.2 Gesamtmortalität.....	21
4.2.3 Lebensqualität.....	21
4.3 Schaden.....	24
4.3.1 Strahlen-induzierte Karzinome (Mortalität).....	24
4.3.2 Überdiagnostik.....	24
4.3.3 Falsch-positive Befunde.....	25
4.3.4 Schmerz durch Mammographie.....	25
4.4 Nutzen/Schaden-Bilanz.....	25
5. Ergebnisse Frauen > 70 Jahre.....	27
5.1 Evidenzgrundlage.....	27
5.1.1 Ergebnis der Literaturrecherche.....	27
5.2 Nutzen.....	27
5.2.1 Brustkrebsmortalität.....	27
5.3 Schaden.....	28
5.3.1 Strahlen-induzierte Karzinome (Mortalität).....	28
5.3.2 Überdiagnostik.....	28
5.3.3 Falsch-positive Befunde.....	28
5.3.4 Schmerz durch Mammographie.....	29
6. Fazit.....	30
7. Literatur.....	31

Abbildungsverzeichnis

Abbildung 1	Nutzenbewertung Mammographie-Screening/Grafik	15
-------------	---	----

Tabellenverzeichnis

Tabelle 1	Endpunkte	18
Tabelle 2	RCTs Mammographie-Screening	20
Tabelle 3	Ergebnisse SR / MA und UK Age Trial ; Endpunkt Brustkrebsmortalität (40-49J) Mammographie-Screening bei Frauen zwischen 40 und 49 Jahren / nach [51]; modifiziert und ergänzt, abgeglichen mit [10]	22
Tabelle 4	Nutzen/Schaden-Bilanz – Frauen 40-49 Jahre	26
Tabelle 5	Ergebnisse RCTs und MA; Endpunkt Brustkrebsmortalität (> 70J)	27
Tabelle 6	Nutzen/Schaden-Bilanz – Frauen > 70 Jahre	29

1. Hintergrund

1.1 Screening

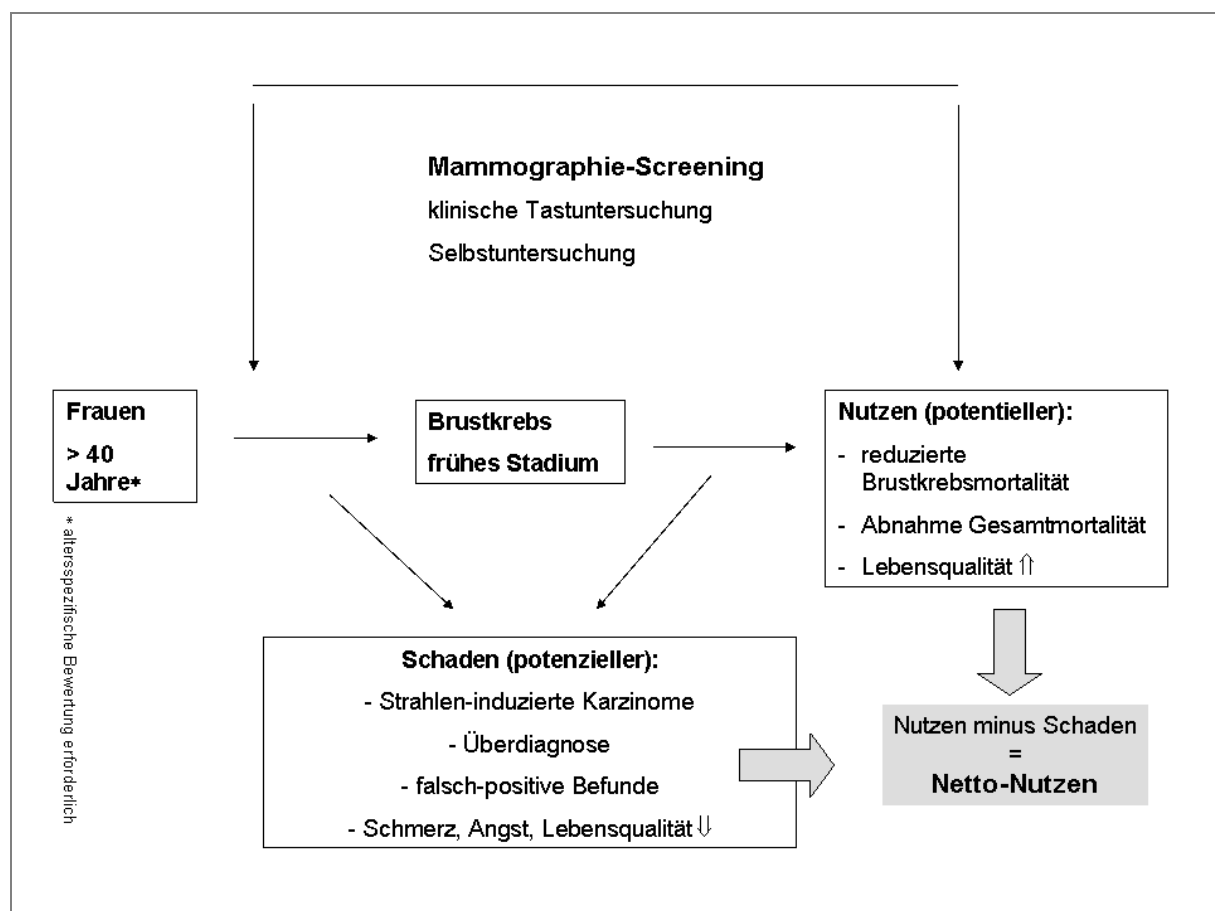
Der im vorliegenden Bericht verwendete Begriff des „Screenings“ ist wie folgt definiert:

Definition:

Unter Screening wird ein Programm verstanden, das durch Ansprache aller Personen einer definierten Zielgruppe (Bevölkerungsbezug) und Vorgaben zur Qualitätssicherung von der ersten Einladung bis zur Bewertung der Endpunkte gekennzeichnet ist. Screening-Programme wenden sich ausschließlich an Personen, die hinsichtlich der Zielerkrankung asymptomatisch sind und sich subjektiv gesund fühlen. Nicht zur Zielgruppe gehören damit Patienten mit Symptomen oder klinischer Manifestation einer Erkrankung. Die Initiative zur Untersuchung geht nicht von den Betroffenen oder deren betreuenden Ärzten aus (modifiziert nach [43]).

Abbildung 1

Nutzenbewertung Mammographie-Screening/Grafik modifiziert nach USPSTF [51]



Es erfolgt eine Beschränkung auf die Beurteilung des Nutzens eines Mammographie-Screenings. Inwieweit ein Screening-Programm eingebettet in multiple Maßnahmen der Brustkrebsfrüherkennung noch einen darüber hinaus reichenden Nutzen haben kann, ist nicht Gegenstand der Bewertung. Ebenso wenig erfolgt hier eine Bewertung für Frauen der erwähnten Altersgruppe, die bereits Symptome aufweisen (kein

Screening) oder die aus eigenem Anlass oder durch Anraten ihres betreuenden Arztes um eine Mammographie ersuchen (Selbsteinladung/ „opportunistisches“ bzw. „graues Screening“).

1.2 Nutzen

Unter Nutzen werden im vorliegenden Bericht zunächst nur die positiven Effekte der Intervention verstanden, die dann den negativen Effekten (Schaden) gegenübergestellt werden. Am Ende resultiert eine Nutzen/Schaden-Bilanz, deren Ergebnis auch als Netto-Nutzen bezeichnet werden kann.

2. Fragestellung

1. Besteht ein nachgewiesener Nutzen eines Mammographie-Screenings für Frauen im Alter von 40-49 Jahren?
2. Besteht ein nachgewiesener Nutzen eines Mammographie-Screenings für Frauen älter als 70 Jahre?

3. Methodik

3.1 Kriterien für den Einschluss von Studien

3.1.1 Population

Asymptomatische Frauen im Alter von 40-49 Jahren und über 70 Jahre.

3.1.2 Intervention und Vergleich

Die Einladung zum Screening ist die untersuchte Intervention. Der Begriff Screening wird dabei entsprechend der einleitend dargestellten Definition verstanden. Als Vergleich werden Frauen herangezogen, die keine Einladung zu einem Screening-Programm erhalten haben. Die „Intervention Einladung zum Screening“ und nicht die „Teilnahme am Screening“ zu untersuchen entspricht dem „intention-to-screen“ Prinzip. In den zum Thema durchgeführten Studien erhält auch ein gewisser Prozentsatz von Frauen in den Kontrollgruppen eine Mammographie („Kontamination“ der Kontrollgruppe). Diese Mammographien werden bei der Auswertung ignoriert und die Frauen werden als Teilnehmerinnen in der Kontrollgruppe (ohne Intervention) ausgewertet, auch wenn sie außerhalb des Programms mammographiert wurden. Diese Auswertung entspricht ebenfalls dem „intention-to screen“ Prinzip.

Hinsichtlich einer möglichen systematischen Verzerrung resultiert bei der angegebenen Vorgehensweise eher eine Unterschätzung bzw. konservative Schätzung des Interventionseffektes. Darüber hinaus entspricht das Vorgehen am ehesten den realen Bedingungen und gewährleistet somit die Übertragbarkeit der Ergebnisse.

3.1.3 Studientypen

Nutzen

Für die Nutzenbewertung einer Intervention gelten randomisierte kontrollierte Studien (RCTs) und deren „gepoolte“ Auswertungen (systematische Übersichtsarbeiten bzw. Meta-Analysen) als die Evidenz der höchsten Stufe in der Beantwortung der Frage, so auch für die Nutzenbewertung eines Mammographie-Screenings. Das bedeutet, dass davon auszugehen ist, dass die in diesen Studien erzielten Ergebnisse mit dem geringsten Risiko einer systematischen Verzerrung einhergehen. Inwieweit sich die Ergebnisse auf den Gültigkeitsbereich der Leitlinie anwenden lassen und welche Gründe möglicherweise die Anwendbarkeit einschränken, wird Teil der Diskussion der Ergebnisse sein.

Betrachtet werden bevorzugt RCTs, die Frauen der genannten Altersgruppen einschließen. Aber auch gepoolte Subgruppenanalysen für Frauen der Altersklassen aus den Daten der vorliegenden RCTs werden zur Beantwortung der Fragen herangezogen. Die aus RCTs, Systematischen Übersichtsarbeiten (SR) und ggf. Meta-Analysen (MA) zur Verfügung stehenden Daten werden dargestellt. Die Darstellung der Ergebnisse aller Meta-Analysen für Frauen der interessierenden Altersklassen dient der Darlegung der Spannweite (Unsicherheit) des möglichen Nutzens und trägt der in der Vergangenheit geführten intensiven kontroversen Diskussion um die Qualität und Validität der zugrunde liegenden RCTs und der darauf aufbauenden Systematischen Übersichtsarbeiten Rechnung [8;17].

Schaden

Zur Beurteilung möglicher Schadwirkungen des Screenings werden neben den RCTs auch Ergebnisse aus Beobachtungsstudien bzw. Registerdaten berücksichtigt. Die im Vergleich größere Unsicherheit hinsichtlich der Validität der Ergebnisse wird dabei beschrieben und in der abschließenden Gegenüberstellung der positiven und negativen Effekte berücksichtigt und dargelegt.

3.2 Literaturrecherche

Nutzen

Die Literaturrecherche erfolgte entsprechend dem Auftrag ab dem 01.01.2000.

Letztes Recherchedatum ist der 15.04.2007.

Die Recherche erfolgte in Medline und der Cochrane Library.

Es wurden folgende Suchbegriffe (in unterschiedlichen Kombinationen) verwendet und mit diesen verschiedene Suchstrategien gebildet.

"Mass Screening"[MeSH]

"Breast Neoplasms"[MeSH] OR "Breast Neoplasms/prevention and control"[MeSH]

"Mammography"[MeSH]

Die Einschränkung hinsichtlich der Studientypen (RCTs und SR) erfolgte über die von der Arbeitsgruppe von Haynes angegebenen Suchfilter, wobei immer die sensitive Suchstrategie verwendet wurde [55].

Zusätzlich wurden noch die Literaturverzeichnisse der gefundenen Arbeiten hinsichtlich weiterer Literaturstellen gescreent, die Suche erbrachte keine zusätzlichen Treffer.

Schaden

Hinsichtlich möglicher Schadwirkungen erfolgte keine systematische Recherche. Als Ausgangspunkt diente die 2007 erschienene Systematische Übersichtsarbeit des American College of Physicians (ACP) in welcher eine explizite Gegenüberstellung von positiven und negativen Effekten eines Mammographie-Screenings enthalten ist [14]. Auf dieser Arbeit beruhen auch die Empfehlungen der ACP Leitlinie zum Thema Mammographie-Screening für Frauen im Alter von 40-49 Jahren aus dem Jahr 2007 [15].

Dem genannten Review sind Evidenztabellen zu folgenden „negativen“ Endpunkten zu entnehmen:

- Strahlenschäden (Mortalität und Inzidenz von Karzinomerkrankungen)
- Überdiagnose
- Falsch-positive Befunde
- Schmerz durch Mammographie

In der Beurteilung bezogen wir uns auf diese Evidenztabellen.

3.3 Endpunkte

Für die vorzunehmende Nutzen/Schaden-Bilanz eines Mammographie-Screenings werden die in Tabelle 1 aufgeführten Endpunkte (orientiert an [14;15]) berücksichtigt.

In der Tabelle enthalten sind Erklärungen hinsichtlich der Bedeutung und der methodischen Validität der berücksichtigten Endpunkte.

Tabelle 1
Endpunkte

positive Effekte/Nutzen durch Screening - Benefit -	Kommentar
Brustkrebsmortalität	Üblicher Endpunkt. Kritik wird geäußert hinsichtlich der Verlässlichkeit der Erhebung, insbesondere hinsichtlich des Übersehens einer therapiebedingten erhöhten Mortalität (z.B. kardiovaskuläre Toxizität). Zunehmende Bemühungen, bereits bei der Durchführung der Studien mit entsprechenden Maßnahmen vorzubeugen (so sollen z.B. Personen, die die Todesursachen kodieren, verblindet sein bezüglich des Screening-Status). Auch bei der Auswertung wird versucht, mit entsprechenden Maßnahmen (Sensitivitätsanalysen) Effekte möglicher Verzerrungen zu ermitteln bzw. zu vermindern.
Gesamtmortalität	Wünschenswerter Endpunkt, der nur in einem Teil der zugrunde liegenden Originalstudien berichtet wird. Da die Intervention nur auf einen geringen Teil des Todesursachenspektrums Einfluss nimmt (4% aller Frauen starben in Deutschland im Jahr 2005 an Brustkrebs, 96% an anderen Todesursachen) ist ein statistisch signifikanter Effekt für diesen Endpunkt kaum zu erwarten. Außerdem müssen andere Todesursachen an die Stelle des verhinderten Brustkrebstodes treten („konkurrierende Mortalität“), wenn auch in einem späteren Lebensalter, so dass die Verminderung eines Brustkrebssterbefalls immer einen zusätzlichen (späteren) Tod bei einer anderen Todesursache nach sich ziehen muss. Eine Untersuchung zum Nachweis statistisch signifikanter Unterschiede mit ausreichend großen Studienkollektiven erscheint unrealistisch.
Verbesserung der Lebensqualität	Bedeutender Endpunkt, der in den zugrunde liegenden RCTs nicht untersucht wurde. Meist werden Surrogatparameter gemessen z.B. Tumor-Stadienschift hin zu kleineren Tumoren, eine weniger eingreifende Therapie, aber auch Besserung der Befindlichkeit nach Erhalt eines unauffälligen Screening-Befundes [39]. Auf die zusammenfassende Darstellung der Ergebnisse hinsichtlich dieses Endpunktes wird auf Grund unzureichender Datenlage verzichtet.

negative Effekte/Schaden durch Screening - Harm -

negative Effekte/Schaden durch Screening - Harm -	Kommentar
Strahleninduzierte Karzinome	In RCTs zum Thema nicht untersucht. Es liegt nur indirekte Evidenz aus anderen Studien vor. Die Beurteilung ist insgesamt mit sehr großer Unsicherheit behaftet.
Überdiagnose/Übertherapie	Hierbei handelt es sich um die Diagnose eines vorliegenden invasiven Karzinoms oder eines in-situ-Karzinoms (insbes. eines duktales in-situ-Karzinoms; DCIS) durch die „Screening-Mammographie“, welches ohne Screening nie symptomatisch geworden wäre und damit weder die Mortalität noch die Lebensqualität der betroffenen Frau beeinflusst hätte. Da diese Karzinome genauso therapiert werden wie außerhalb des Screenings entdeckte, führt eine Überdiagnose auch zur Übertherapie. Es liegt nur indirekte Evidenz aus anderen Studien vor. Die Beurteilung ist insgesamt mit sehr großer Unsicherheit behaftet [24;25].
Beeinträchtigung der Lebensqualität durch vorverlegte Diagnose	In der Regel wird durch die Screening-Mammographie ein Brustkrebs 3-4 Jahre früher diagnostiziert als ohne Intervention. Das Wissen über eine bösartige Erkrankung kann als Minderung der Lebensqualität empfunden werden. Auf die zusammenfassende Darstellung der Ergebnisse hinsichtlich dieses Endpunktes wird aufgrund unzureichender Datenlage verzichtet.

negative Effekte/Schaden durch Screening - Harm -	Kommentar
falsch-positive Befunde	Fälschlicherweise diagnostizierte Karzinome (invasiv und in-situ), obwohl tatsächlich kein Karzinom vorhanden ist. Die Angaben beziehen sich auf die Mammographie, zu den Unsicherheiten der pathologischen Diagnostik siehe dort. Zu unterschieden sind kumulative Raten und Raten bei jeder neuen Mammographie. Es bestehen große Unterschiede hinsichtlich länderspezifischer Ergebnisse (Nordamerika versus Europa). Die Unterschiede sind durch einen differenten soziokulturellen Hintergrund, insbesondere bei Fragen zur ärztlichen Haftpflicht bei nicht entdeckten Tumoren bedingt. Ein geringerer Anteil an falsch-positiven Befunden wird in der Regel erkaufte mit einem höheren Anteil an Intervallkarzinomen. Die zusammenfassende Beurteilung der Ergebnisse ist mit Unsicherheit behaftet.
falsch-negative Befunde/ Intervallkarzinome	Kein Screening-Test weist eine Sensitivität von 100% auf. Vorhandene Karzinome, die vom Screening nicht identifiziert werden, können entweder wirklich nicht mit dem Screening-Testinstrument erkannt worden sein oder sie sind übersehen worden (falsch-negativer Befund). Um die Sensitivität der Mammographie zu erhöhen, wurde die Doppelbefundung der Aufnahmen eingeführt; Resultate siehe dort. Ist ein im Intervall zwischen zwei Screening-Runden auftretendes Karzinom auch retrospektiv unter Hinzuziehung aller vorhandenen Mammogramme zum Zeitpunkt der letzten Screening-Untersuchung nicht zu erkennen, ist von einem echten Intervall-Karzinom auszugehen und nicht von einem falsch-negativen Befund.
Schmerz durch Mammographie	Bisher wenig systematisch untersucht, große Unsicherheit bezüglich der Präzision der Angaben. Auf die zusammenfassende Darstellung der Ergebnisse hinsichtlich dieses Endpunktes wird aufgrund unzureichender Datenlage verzichtet.

3.3.1 Nutzen/Schaden –Bilanz

Zusammenfassend wird am Ende des Berichts versucht, eine Nutzen/Schaden-Bilanz anzugeben. Die gewählte Darstellung beruht auf der Methodik der GRADE Arbeitsgruppe (Grading of Recommendations, Assessment, Development and Evaluation) [56].

Nach dieser Methodik wird die einer Empfehlung zugrunde liegende Evidenz hinsichtlich ihrer methodischen Qualität einerseits und der resultierenden Effektstärke andererseits ausgewertet. Beides bildet die Grundlage für eine abwägende Entscheidung hinsichtlich des aus positiven und negativen Effekten resultierenden Netto-Nutzens. Diese Darlegungen bilden die Grundlage der Empfehlungsformulierung. Besonderer Wert wird dabei auf die getrennte Bewertung der Datenlage hinsichtlich der relevanten Endpunkte gelegt, welche eine explizite Nutzen/Risiko-Abwägung ermöglicht [49;50].

Nach GRADE erfolgt eine Hierarchisierung der berücksichtigten Endpunkte auf einer Skala von 1-9, über die innerhalb der Leitliniengruppe eine Konsensentscheidung herbeigeführt werden sollte. Diese Hierarchisierung haben wir, da unsere Aufgabe in der Darlegung und Bewertung der vorliegenden Evidenz bestand, nicht durchgeführt. Wir haben somit auf eine Gewichtung der relevanten Endpunkte verzichtet. Das Vorgehen erscheint gerechtfertigt, da unsererseits keine Empfehlung formuliert wird.

4. Ergebnisse Frauen 40-49 Jahre

4.1 Evidenzgrundlage

Tabelle 2 zeigt alle randomisiert-kontrollierten Studien, welche die Wirksamkeit eines Mammographie-Screenings im Hinblick auf die Senkung der Gesamt- und Brustkrebs bedingten Mortalität untersuchen. Bis 2006 konnten auch Angaben zu Frauen im Alter von 40-49 Jahren nur diesen Studien entnommen werden.

Tabelle 2
RCTs Mammographie-Screening

	Publikation	Studienbeginn	Anzahl Frauen	Alter bei Einschluss	Screening Intervall (Monate)	Anzahl "Runden"	Kommentar
1 [57] ¹	HIP / New York	1963	60.495	40-64	12	4	Mammographie + CBE ²
2 [65]	Two County Trial / Kopparberg und Ostergötland	1977	133.065	40-74	24-33	3	Mammographie einzige Studie, die Frauen > 70 Jahre untersucht
3 [64]	Malmö	1976-78	42.283	45-70	18-24	9	Mammographie
4 [58]	Kanada 1 / CNBSS-1	1980	50.430	40-49	12	4-5	Mammographie + CBE
5 [59;60]	Kanada 2 / CNBSS-2	1980	39.405	50-59	12	4-5	Mammographie + CBE
6 [61]	Edinburgh	1978	54.643	45-64	24	4	Mammographie + CBE
7 [62]	Göteborg	1982	49.533	39-59	18	5	Mammographie
8 [63]	Stockholm	1981	60.261	40-64	24-28	2	Mammographie
9	Age Trial Moss et al. 2006 [12]	1991	160.921	39-41	12	4,5 bzw. 5,6 (für Frauen mit ≥ 1 Teilnahme)	

Für eine grundsätzliche Qualitätsbeurteilung der bis 2002 erschienen RCTs (1-8) zum Thema Mammographie-Screening verweisen wir auf die erste Auflage der Leitlinie (Stufe-3-Leitlinie-Brustkrebsfrüherkennung in Deutschland). Auf die Darlegung der kontroversen internationalen Diskussion wird verzichtet, da sie für die Beantwortung der Frage nicht zielführend ist. Eine detaillierte Darlegung des „Age Trials“ [12] kann dem Anhang entnommen werden. Hierbei handelt es sich um die bisher einzige randomisierte kontrollierte Studie, in welche Frauen im Alter von 39 bis 41 Jahren eingeschlossen wurden, für die der mögliche Nutzen eines Screenings – bei einer mittleren Nachbeobachtung von 10,7 Jahren – also im Alter bis 50 Jahre auftritt und nicht in der Altersgruppe der über 50jährigen [29].

¹ Orientiert an USPSTF [51] wurde immer nur eine Literaturstelle für die Studien angegeben und auf die vollständige Angaben aller Publikationen (20-50) pro RCT verzichtet.

² CBE = professionelle Palpation der Brust (clinical breast examination)

4.1.1 Ergebnis der Literaturrecherche

Nach Elimination von Dubletten und Doppelpublikationen blieben 3 SR bzw. MA und 1 RCT, die nach 2000 erschienen sind und zur Bewertung des Nutzens eines Mammographie-Screenings verwendet werden konnten.

Als Ausgangspunkt für die Darstellung des Nutzens für die Gruppe der Frauen dient der Evidenzbericht der US Preventive Services Task Force (USPSTF) aus dem Jahr 2002, welcher eine Aufbereitung der Evidenz für Frauen im Alter von 40-49 Jahren enthält.

Tabelle 3 enthält alle Meta-Analysen, welche (unter anderem) die Frage des Nutzens eines Mammographie-Screenings für Frauen im Alter zwischen 40 und 49 Jahren untersuchen.

4.2 Nutzen

4.2.1 Brustkrebsmortalität

Mit Ausnahme der ersten Systematischen Übersichtsarbeit von Gotzsche und Olsen [42], die aus methodischen Gründen lediglich zwei der vorliegenden RCTs in die Auswertung einbezogen, kommen alle anderen systematischen Übersichtsarbeiten und Meta-Analysen trotz Variabilität hinsichtlich der eingeschlossenen Studien und der angewendeten statistischen Verfahren zu weitgehend übereinstimmenden Ergebnissen hinsichtlich des Effektmaßes. Die Spanne des ermittelten Relativen Risikos der Brustkrebsmortalität beträgt 0,77-0,96 unter Mammographie-Screening gegenüber Nicht-Screening. In etwa der Hälfte der MA wird dabei das 95%-Signifikanzniveau gerade erreicht, in der anderen Hälfte knapp verfehlt.

In einer neueren systematischen Übersichtsarbeit kommt auch Gotzsche zu einer Reduktion des Relativen Risikos der Brustkrebsmortalität, die bei einer mittleren Nachbeobachtungszeit von 13 Jahren 0,84 (95% KI 0,72-0,99) beträgt. Dieses Ergebnis stimmt weitgehend überein mit dem der Meta-Analyse der USPSTF von 2002, die ein Relatives Risiko von 0,85 (95% KI 0,73-0,99) bei einer mittleren Nachbeobachtungszeit von 14 Jahren ermitteln.

Die Ergebnisse der aufgeführten Subgruppenanalysen werden durch die des „AGE Trials“ bestätigt. Das ermittelte Relative Risiko beträgt 0,83 (95% KI 0,66-1,04), bei einer mittleren Nachbeobachtungszeit von 10,7 Jahren. Die absolute Risikoreduktion beträgt 0,037 pro 1.000 Frauenjahre. Das Screening-Intervall ist 12 Monate. Die mittlere Anzahl an Screening-Runden ist 4,5 bzw. 5,6 für die Frauen, die mindestens einmal teilnahmen;

81% der Frauen nahmen an wenigstens einer Mammographie teil. Aus der Risikoreduktion würde sich (Achtung Ergebnis nicht statistisch signifikant) eine errechnete Number Needed to Screen von 2.512 errechnen, um einen Todesfall aufgrund von Brustkrebs in 10 Jahren zu verhindern. Im Mittel erhielten die Frauen in dieser Zeit sieben Einladungen zum Mammographie-Screening.

4.2.2 Gesamtmortalität

Siehe hierzu auch Tabelle 1. Valide Angaben sind aufgrund der spärlichen Datenlage nicht möglich.

Nyström gibt für Frauen der Altersgruppe 40-44 Jahre ein Relatives Risiko von 1,03 (95% KI 0,92-1,15) an, für Frauen der Altersgruppe 45-49 Jahre von 0,99 (95% KI 0,89-1,10) und für den 10-Jahreszeitraum von 1,00 (95% KI 0,95-1,06) [52]. Somit lässt sich in keiner der drei angegebenen Gruppen ein Effekt durch Screening in Bezug auf die Gesamtmortalität zeigen.

4.2.3 Lebensqualität

Siehe hierzu auch Tabelle 1. Valide Angaben sind aufgrund der spärlichen Datenlage nicht möglich.

Tabelle 3

Ergebnisse SR / MA und UK Age Trial ; Endpunkt Brustkrebsmortalität (40-49J)

Mammographie-Screening bei Frauen zwischen 40 und 49 Jahren / nach [51]; modifiziert und ergänzt, abgeglichen mit [10]

Publikation / Systematische Übersichtsarbeit oder Meta-Analyse	Qualitätsbewertung der eingeschlossenen Studien	Eingeschlossene Studien	Statistisches Modell	Mittlere Nachbeobachtung in Jahren	Brustkrebsmortalität Relatives Risiko (95% KI ³)	Number Needed to Screen (bezogen auf eingeladene Frauen und Modi über 10 Jahre nach Studie, Angabe in Publikation)
Larsson et al., 1997 [20]	nein	Malmö; Two County; Göteborg; Stockholm	Weighted relative risks	12,8	0,77 (0,59-1,01)	keine
Nystrom et al., 1993 [19]	ja, IPD	Malmö; Osetrgötland aus Two County; Göteborg; Stockholm	RR durch density Methode; 2 Modelle zu Grunde legend: follow-up und evaluation		Evaluation-Model 0,8 (0,63-1,01) Follow-up-Model 0,91 (0,76-1,09)	keine
Cox, 1997 [21]	nein	HIP; Malmö; Two County; Kanada-1; Edinburgh; Göteborg; Stockholm	Fixed effects	10	0,93 (0,77-1,11)	keine
Glasziou and Irwig, 1997 [3;6]	ja, alle Studien wurden mit "gut" bewertet, Malmö und Kanada-1 am besten, Two County und Göteborg am schlechtesten	HIP; Malmö; Two County; Kanada-1; Edinburgh; Göteborg; Stockholm	Variance-weighted	13,13	0,85 (0,71-1,01)	keine
Hendrick et al., 1997 [7]; Aktualisierung von Smart et al., 1995 [1]	nein	HIP; Malmö; Two County; Kanada-1; Edinburgh; Göteborg; Stockholm	Fixed effects	12,7	0,82 (0,71-0,95)	keine

³ KI = Konfidenzintervall

Publikation / Systematische Übersichtsarbeit oder Meta-Analyse	Qualitätsbewertung der eingeschlossenen Studien	Eingeschlossene Studien	Statistisches Modell	Mittlere Nachbeobachtung in Jahren	Brustkrebsmortalität Relatives Risiko (95% KI³)	Number Needed to Screen (bezogen auf eingeladene Frauen und Modi über 10 Jahre nach Studie, Angabe in Publikation)
Kerlikowske et al., 1995 [5]	nein	HIP; Malmö; Two County; Kanada-1; Edinburgh; Göteborg; Stockholm	Fixed effects	10-12	0,83 (0,65-1,06)	keine
Berry, 1998 [22]	nein	HIP; Malmö; Two County; Kanada-1; Edinburgh;	Random effects	12-15	0,82 (0,49-1,17)	keine
Olsen and Gotzsche, 2001 [42]	ja, nur Malmö und CNBSS eingeschlossen, alle anderen Studien aus methodischen Gründen ausgeschlossen	Kanada-2, Malmö	Fixed effects	13	1,03 (0,77-1,38)	keine
USPSTF, 2002 [51]	ja, Edinburgh ausgeschlossen, alle anderen Studien einbezogen	HIP; Malmö; Two County; Kanada-1; Göteborg; Stockholm	Random effects	ca. 14	0,85 (0,73-0,99)	1.698
Gotzsche and Nielsen, 2006 [9]	ja, Edinburgh ausgeschlossen	HIP; Malmö; Two County; Kanada-2; Göteborg; Stockholm	Je 2 Auswertungen Sensitivitätsanalyse in Abhängigkeit von Studienqualität; Ergebnis hier für alle Studien Fixed effect	für 7 J und 13 J 7 J ----- 13 J	0,96 (0,78-1,18) ----- 0,84 (0,72-0,99)	keine
Moss et al., 2006 [12] Meta-Analyse	unklar	HIP; Malmö; Two County; Kanada-2; Edinburgh; Göteborg; Stockholm; Age Trial	Keine Angabe	ca. 11	0,84 (0,74-0,95)	keine
Moss et al., 2006 [12], RCT	nicht anwendbar	nicht anwendbar	nicht anwendbar	10,7 Jahre	0,83 (0,66-1,04)	2.512

4.3 Schaden

4.3.1 Strahlen-induzierte Karzinome (Mortalität)

Siehe hierzu auch Tabelle 1.

Für die Beantwortung dieser Frage liegt nur indirekte Evidenz vor, da entsprechende Studien mit der bei der Mammographie verwendeten Strahlenart und -dosis fehlen.

Die Angaben zur Strahlenbelastung schwanken in der Literatur zwischen 4-5 mGy pro 2-Ebenen-Mammographie [11;14;23]. Es wird angenommen, dass die kumulative Organdosis für Frauen mit 50 Jahren, die mit 40 Jahren in ein Mammographie-Screening-Programm aufgenommen wurden und eine jährliche Mammographie erhielten, bei ca. 60 mGy liegt [14].

Basierend auf einer statistischen Modellbildung (Grundlage sind Daten aus dem Screening-Programm des National Health Service in Großbritannien – in das Modell gehen unvalidierte Annahmen ein) gibt es Angaben, die von einer Mortalitätsrate aufgrund strahleninduzierten Brustkrebses von 0,22-0,5 / pro 1.000 Frauen bei 10 Jahren Screening, Beginn mit 40 Jahren und jährlicher Mammographie, ausgehen. Das würde einer Zahl von NNH (Number Needed to Harm) von 2.000-4.545 in 10 Jahren entsprechen [45;36]. Diese Zahl liegt in vergleichbarer Größenordnung wie der im Age Trial errechnete Nutzen von NNS= 2.512. In der Netto-Bilanz wäre als pro verhindertem Brustkrebs-Sterbefall mit einer ähnlich hohen Zahl Strahlen-induzierter zusätzlicher Mammakarzinome zu rechnen, die allerdings in der Mehrzahl nicht zu einem Brustkrebs-Sterbefall führen.

Eine britische Arbeitsgruppe kommt auf der Basis mathematischer Modelle zu dem Ergebnis, dass bei Einschluss von Frauen unter 40 Jahren in ein Screening-Programm ein Netto-Schaden resultiert und erst ab dem Alter von 40 Jahren ein Netto-Nutzen zu erwarten ist [11].

Alle Angaben sind mit einer großen Unsicherheit behaftet, da alle Berechnungen einer Vielzahl von Annahmen unterliegen und die Validität entsprechend eingeschränkt ist. So ist umstritten, ob Daten aus dem Hoch-Dosisbereich (Atombombenversuche oder Strahlentherapie) linear auf niedrige Dosen in der Röntgendiagnostik umgerechnet werden können, ob es eine Schwellendosis ohne Risiko gibt („linear no threshold- Hypothese“) und ob Frauen mit einem (familiär) erhöhten Brustkrebsrisiko empfänglicher sind für Strahlen-induzierte Karzinome [44].

Eine Zusammenstellung der indirekten Evidenz findet sich in Tabelle 2 des Appendix der Systematischen Übersichtsarbeit für das ACP [14].

4.3.2 Überdiagnostik

Siehe hierzu auch Tabelle 1.

Für Frauen der Altersgruppen < 50 Jahre gibt es keine Aussagen zur Häufigkeit von Überdiagnosen, so dass nur eine Übertragung der Zahlen aus einer anderen Altersgruppe möglich ist.

Literaturangaben zur zusätzlichen Inzidenz (Überdiagnose, korrigiert für die zeitliche Vorverlagerung der Diagnose [lead time]) des invasiven Brustkrebses beim Mammographie-Screening (Frauen im Alter 50-69 Jahre) schwanken zwischen erwarteten 2% in den Niederlanden [46] und beobachteten 3% in Italien [48] bis zu 45 % in Norwegen bzw. über 50% in Schweden [33]. Bezieht man auch in-situ-Karzinome mit in die zusätzliche Inzidenz ein, beträgt der Aufschlag insgesamt zwischen 5% [48] und 80% [33]. Die anderen zitierten Literaturstellen weisen Prozentsätze zwischen den genannten Extremwerten aus [18; 42; 26, 47]⁴.

Screening erhöht die Wahrscheinlichkeit einer DCIS-Diagnose. Der Anteil der in-situ-Karzinome an allen inzidenten Mammakarzinomen steigt durch ein Screening-Programm [40;41]. Dabei ist die Frage, welcher Anteil von DCIS sich nicht zu einem invasiven Karzinom entwickelt und damit auch weder Lebenserwartung noch Lebensqualität der Frau beeinflusst, ungeklärt. Aus älteren Fallserien lässt sich rückschließen, dass sich DCIS in einem Prozentsatz zwischen 14% und 60% zu einem invasiven Karzinom entwickeln. Allerdings ist die Datenqualität für diese Aussage sehr schlecht. Es liegt eine größere Fallserie (n= 1.063) vor; hier entwickelte sich in 8,2% aller Fälle ein invasives Karzinom, allerdings wurden die Frauen behandelt.

Siehe auch Tabelle 3 des Appendix der Systematischen Übersichtsarbeit für das ACP [14].

⁴ Die Literatur zur Überdiagnose wurde zusätzlich an Hand der aktuellen Arbeit von Zackrisson [18] und den dazu eingegangenen Leserbriefen und der in diesen zitierten Literatur abgeglichen.

4.3.3 Falsch-positive Befunde

Siehe hierzu auch Tabelle 1.

Bei der Screening-Mammographie in der Altersgruppe 50-69 Jahre liegt die Rate an falsch-positiven Befunden zwischen 2% und 4%. Die EU-Leitlinien geben einen Korridor von 3-7% für die Recall-Rate (% Frauen einbestellt zur Abklärung) vor, die sich zusammensetzt aus der Detektionsrate (bestätigtes Karzinom) und der Falsch-Positiv-Rate. Die Häufigkeit schwankt in Abhängigkeit von der angestrebten Sensitivität und Spezifität der Untersuchung und ob es sich um die erste Untersuchung bei einer Frau handelt oder um eine Folgeuntersuchung. Deutliche Unterschiede bestehen zwischen Europa und den USA. In den USA wird - kulturell bedingt - größerer Wert auf eine möglichst hohe Sensitivität der Untersuchung gelegt, so dass hier die Rate an falsch-positiven Befunden deutlich über der Rate in europäischen Ländern liegt. Besonders in den Niederlanden wollte man die Anzahl der falsch-positiven Befunde gering halten, was zu einer relativ hohen Zahl von im Intervall zwischen zwei Screening-Runden diagnostizierten Karzinomen (Intervall-Karzinomen) führte [40].

Insgesamt muss von einer kumulativ steigenden Wahrscheinlichkeit eines falsch-positiven Befundes ausgegangen werden: Zahlen aus den USA zeigen, dass bei Frauen im Alter von 40-49 Jahre die kumulative Rate falsch positiver Befunde bei 38% nach 5 Mammographien und bei 49% nach 10 Mammographien im Alter von 40-69 Jahren liegt [30]. Dieses würde bei einer Gruppe von 10.000 Frauen, die mit 40 Jahren in ein Screening-Programm mit jährlicher Untersuchung einbezogen werden, in einem Zeitraum von 10 Jahren 5.000 zur Abklärung einbestellte Frauen und bei ca. 1.900 Frauen eine Biopsie zur Folge haben [54]. In Norwegen wurden bei 10 Mammogrammen im Zeitverlauf bei 20 % der Frauen ein falsch-positiver Befund erhoben [14; 31; 16].

Die Beeinflussung der Teilnahmerate an der nächsten Mammographie-Runde durch falsch-positive Befunde ist von Land zu Land verschieden. Bei europäischen Frauen ließ sich bisher kein Zusammenhang zeigen [27].

Siehe auch Tabelle 4 des Appendix der Systematischen Übersichtsarbeit für das ACP [14].

4.3.4 Schmerz durch Mammographie

Siehe hierzu auch Tabelle 1.

Der Prozentsatz von Frauen, die Schmerzen durch die Mammographie angeben, schwankt zwischen 28% und 77% [14]. Es besteht allerdings keine Korrelation zwischen erlebtem Schmerz und der Bereitschaft, sich an einer nächsten Screening-Untersuchung zu beteiligen [14].

Siehe auch Tabelle 5 des Appendix der Systematischen Übersichtsarbeit für das ACP [14].

Narbenbildung durch Biopsie

Wenn zur diagnostischen Abklärung eine offene Biopsie gewählt wird, ist mit einer kosmetisch und radiologisch sichtbaren Narbenbildung zu rechnen, die gegebenenfalls die Befundung bei der folgenden Screening-Untersuchung beeinträchtigt. Werden statt offener Biopsien Vakuum-Stanzbiopsien vorgenommen, soll eine Narbenbildung oder Beeinträchtigung der radiologischen Befundung bei Folge-Mammographien jedoch weitgehend auszuschließen sein (siehe auch Evidenzbericht zur Frage der Vakuumbiopsien, Tabelle 1).

4.4 Nutzen/Schaden-Bilanz

Die Darstellung beruht auf der oben aufgeführten Methodik der GRADE Arbeitsgruppe (Grading of Recommendations, Assessment, Development and Evaluation).

Die Datenlage für Frauen im Alter zwischen 40 und 49 ist hinsichtlich der Anzahl der in methodisch validen Studien untersuchten Frauen (>160.000) überwältigend. Es kann als unwahrscheinlich bezeichnet werden, dass weitere Studienergebnisse die ermittelten Effekte wesentlich beeinflussen werden. Der relative Nutzen beträgt 15%. Die Erkrankung ist allerdings so selten, dass das Resultat nicht statistisch signifikant ist. Die resultierende absolute Risikoreduktion ist gering, 2.500 Frauen müssten über 10 Jahre zum Mammographie-Screening eingeladen werden, damit bei einer Frau der Tod durch Brustkrebs verhindert werden kann.

Sehr viel unsicherer ist die Datenlage hinsichtlich negativer Effekte insbesondere Strahlen-induzierte Karzinome, weshalb die Bewertung des so genannten Netto-Nutzens deutlich schwerer fällt. Dies gilt umso mehr, da nicht auszuschließen ist, dass der potentielle Schaden den Nutzen aufwiegen könnte. Überdiagnosen und falsch positive Befunde sind in der Bilanz ebenfalls zu berücksichtigen.

Tabelle 4
Nutzen/Schaden-Bilanz – Frauen 40-49 Jahre

	Endpunkte	Datenqualität	Effektgröße
Nutzen	Brustkrebsmortalität	⊕ ⊕ ⊕ ⊕	RR . 0,84 (0,74-0,95) NNS ca. 2500 ⁵
	Gesamtmortalität	keine ausreichenden Daten	
	Verbesserung der Lebensqualität	keine ausreichenden Daten	
Schaden	Strahlen-induzierte Karzinome (Mortalität)	⊕	NNH ca. 2.500-4.500 Die Zahlenangaben sind mit einer großen Unsicherheit nicht nur hinsichtlich der Validität der Daten, sondern auch hinsichtlich der Größe des Effektes behaftet.
	Überdiagnose	⊕	2%-50% höhere Inzidenz (ohne in situ Karzinome) ----- 5%-80% höhere Inzidenz (inklusive in situ Karzinomen)
	Falsch-positive Befunde	⊕ ⊕	20 – 49% (kumulativ nach 10 Mammographien)
	Schmerz durch Mammographie	keine ausreichenden Daten	

RR = Relatives Risiko

NNH = Numbers Needed to Harm

- ⊕ ⊕ ⊕ ⊕ Eine **hohe Evidenzstufe** bedeutet: Es ist **unwahrscheinlich**, dass weitere Forschung in diesem Bereich unser Vertrauen in den geschätzten Effekt ändern wird.
- ⊕ ⊕ ⊕ Eine **mittlere Evidenzstufe** bedeutet: Es ist **wahrscheinlich**, dass unser Vertrauen in den geschätzten Effekt durch weitere Forschung erheblich beeinflusst wird und sich der geschätzte Effekt dadurch ändern kann.
- ⊕ ⊕ Eine **niedrige Evidenzstufe** bedeutet: Es ist **sehr wahrscheinlich**, dass unser Vertrauen in den geschätzten Effekt durch weitere Forschung erheblich beeinflusst wird und sich der **geschätzte Effekt** dadurch wahrscheinlich **ändern wird**.
- ⊕ Eine **sehr niedrige Evidenzstufe** bedeutet: Jede Effektschätzung ist **sehr unsicher**.

⁵ Anzahl der über 10 Jahre wiederholt einzuladenden Frauen, bis eine Frau bezüglich des Endpunktes profitiert. Nur Angaben einer Größenordnung, da als Quelle MA verwendet.

5. Ergebnisse Frauen > 70 Jahre

5.1 Evidenzgrundlage

Von den in Tabelle 2 aufgeführten randomisierten kontrollierten Studien zum Nutzen eines Mammographie-Screenings untersucht nur die Ostergötland Studie [66] (Teil der Two County Study) eine größere Gruppe Frauen der interessierenden Altersklasse.

Hier wurden insgesamt etwa 10.000 Frauen > 70 Jahre einbezogen. Die einzige weitere randomisierte kontrollierte Studie, die Frauen > 70 Jahre untersucht ist die Malmö -1 Studie [67], allerdings nur ca. 600 Frauen.

5.1.1 Ergebnis der Literaturrecherche

Die Literaturrecherche war identisch mit der für die jüngeren Frauen. Die Auswertung der Literaturverzeichnisse ergab keine zusätzlichen Treffer. Bei Durchsicht der vorhandenen systematischen Übersichtsarbeiten war bereits eine Arbeit identifiziert worden, die eine Auswertung der Daten für Frauen > 70 Jahre vornimmt [52].

5.2 Nutzen

5.2.1 Brustkrebsmortalität

Die einzige Auswertung gepoolter Daten für Frauen dieser Altersgruppe können einer Meta-Analyse der schwedischen Screening Studien entnommen werden [52].

Tabelle 5

Ergebnisse RCTs und MA; Endpunkt Brustkrebsmortalität (> 70J)
Mammographie-Screening bei Frauen älter 70 Jahre / nach [52]

Publikation/ Systematische Übersichtsarbeit oder Meta-Analyse	Eingeschlossene Studien	Anzahl Frauen	Mittlere Nachbeobachtung in Jahren	Brustkrebsmortalität Relatives Risiko (95% KI)
Nyström et al., 2002 [52]	Malmö 1 und Ostergötland	11.478	17,4 bzw. 17,2	Modell 1 1,18 (0,71 – 1,97) ----- Modell 2 1,12 (0,73 – 1,72)

Für Frauen der 10 Jahres-Altersgruppe 65-74 Jahre zeigt sich ein Nutzen durch die Mammographie. Das Relative-Risiko Brustkrebsmortalität beträgt für die Gruppe der 65-74 Jährigen 0,81 (KI 95% 0,61-1,07) bei Auswertung nach dem Evaluationsmodell und 0,78 (KI 95% 0,62-0,99) bei Auswertung nach dem Follow-up-Modell.

Die differenzierte Betrachtung der zugrunde liegenden 5-Jahres-Alterskohorten (65-69 Jahre und 70-74 Jahre) zeigt ein anderes Ergebnis. Für die älteren Frauen (70-74 Jahre) errechnet sich bei Anwendung beider Modelle eine erhöhtes Risiko für Brustkrebs-Mortalität: RR 1,18 (KI 95% 0,71-1,97), bzw. RR 1,12 (KI 95% 0,73-1,72).

In der jüngeren Altersgruppe (65-69 Jahre) ergibt die Auswertung einen deutlichen Nutzen RR 0,69 (KI 95% 0,49-0,96), bzw. RR 0,68 (KI 95% 0,52-0,89). Da etwa doppelt so viele Frauen der jüngeren

Alterskohorte angehören, bestimmt diese Gruppe das Ergebnis für die 10-Jahres-Alterskohorte (65-74 Jahre) und den hier zu vermutenden Nutzen. Die Ergebnisse für die Gesamtgruppe können daher nicht zur Beantwortung der Frage nach dem Effekt eines Screenings bei Frauen über 70 Jahre herangezogen werden.

Soll explizit die Frage nach Frauen im Alter von 70-74 Jahren beantwortet werden, dann ist anhand dieser einzigen vorliegenden Meta-Analyse eine Erhöhung der Brustkrebs-induzierten Mortalität durch Mammographie-Screening wahrscheinlicher als eine Senkung der Mortalität. Allerdings sind die Ergebnisse statistisch nicht signifikant und aufgrund der spärlichen Datenlage mit großer Unsicherheit behaftet.

Den ungünstigen Ergebnissen aus dem RCT stehen Prozessdaten mit günstigen Ergebnissen aus dem niederländischen Screening-Programm gegenüber, die allerdings ohne Vergleichsgruppe gewonnen wurden und daher von geringerer Aussagekraft sind [68]. Zusammenfassend muss festgestellt werden, dass die Aussagen hinsichtlich des potentiellen Nutzens eines Mammographie-Screenings für die Frauen > 70 Jahre mit großer Unsicherheit behaftet sind, was durch die geringe Zahl von Frauen der interessierenden Altersgruppe in den Studien bedingt ist.

5.3 Schaden

5.3.1 Strahlen-induzierte Karzinome (Mortalität)

Siehe hierzu auch Tabelle 1 und Kapitel 4.3.1.

Weitere Studien bzw. getrennte Auswertungen für Frauen > 70 Jahre liegen zum Strahlenrisiko nicht vor. Bei einem Screening-Programm, welches mit 50 Jahren beginnt und über das Alter von 70 Jahren hinaus fortgeführt wird, steigt die kumulative Strahlenbelastung. Es liegen keine verlässlichen Zahlen vor, die eine Abschätzung des Risikos erlauben würden. Wichtig ist anzumerken, dass die Strahlensensibilität des Gewebes mit steigendem Alter abnimmt. Die mindestens 10jährige Latenzzeit bis zur Entwicklung eines Strahlen-induzierten Karzinoms ist bei der Abschätzung der Mortalität durch induzierte Karzinome ebenfalls zu berücksichtigen. Mit zunehmendem Lebensalter fällt die Wahrscheinlichkeit, ein Strahlen-induziertes Karzinom noch zu erleben.

5.3.2 Überdiagnostik

Siehe hierzu auch Tabelle 1 und Kapitel 4.3.2.

Für Frauen im Alter > 70 Jahren liegen keine gesonderten Auswertungen vor. Die Häufigkeit von DCIS sinkt mit steigendem Alter, so dass anzunehmen ist, dass Überdiagnostik für DCIS in höherem Alter eine geringere Bedeutung hat. Dem steht entgegen, dass mit höherem Lebensalter die konkurrierende Mortalität (Brustkrebs nicht Todesursache) eine größere Bedeutung bekommt und die verbleibende Lebenszeit, in der ein noch nicht klinisch auffälliges Mammakarzinom symptomatisch werden kann, abnimmt. Auch ist das Tumorwachstum im höheren Alter verlangsamt, so dass die Wahrscheinlichkeit einer Überdiagnose hierdurch steigt.

5.3.3 Falsch-positive Befunde

Siehe hierzu auch Tabelle 1 und Kapitel 4.3.3.

Für Frauen im Alter > 70 Jahren liegen keine Zahlen aus systematischen Auswertungen der Screening-Programme oder RCTs hinsichtlich der Rate an falsch-positiven Befunden vor. Es müssten also Zahlen aus jüngeren Altersgruppen übertragen werden plus weiterer Aufschläge für zusätzliche Mammographien nach Vollendung des 70. Lebensjahres. Genaue Zahlen sind nicht bekannt.

5.3.4 Schmerz durch Mammographie

Siehe hierzu auch Tabelle 1 und Kapitel 4.3.4.
Angaben für Frauen dieser Altersgruppe liegen nicht vor.

Tabelle 6
Nutzen/Schaden-Bilanz – Frauen > 70 Jahre

	Endpunkte	Datenqualität	Effektgröße
„Nutzen“⁶	Brustkrebsmortalität	⊕⊕⊕ ⁷	RR 1,18 95% KI: 0,71-1,97
	Gesamtmortalität	keine ausreichenden Daten	
	Verbesserung der Lebensqualität	keine ausreichenden Daten	
Schaden	Strahleninduzierte Karzinome (Mortalität)	⊕	Die Zahlenangaben sind mit einer großen Unsicherheit nicht nur hinsichtlich der Validität der Daten sondern auch hinsichtlich der Größe des Effektes behaftet. Zusätzlich zur eingeschränkten methodischen Qualität wurden sie an Frauen anderer Altersgruppen erhoben.
	Überdiagnose	⊕	
	Falsch-positive Befunde	⊕	
	Schmerz durch Mammographie	keine ausreichenden Daten	

Die zusammenfassend dargestellten Resultate für Frauen im Alter über 70 Jahren machen deutlich, dass die Datenlage sowohl hinsichtlich eines potenziellen Nutzens als auch hinsichtlich eines potenziellen Schadens sehr unsicher ist. Aussagekräftige Daten aus Studien liegen nicht vor, so dass eine Nutzenbewertung nicht vorgenommen werden kann. Die zur Verfügung stehenden Daten legen nahe, dass ein Mammographie-Screening bei Frauen über 70 Jahren eher mit einem Schaden als mit einem Nutzen einhergeht.

⁶ Darstellung des Nutzens in Anführungsstrichen, da in RCTs bislang nur ein erhöhtes Relatives Risiko geschätzt wurde, was einem Schaden entsprechen würde.

⁷ Spärliche Datenlage (wenig Frauen der interessierenden Altersgruppe untersucht) führt zu einer Abwertung der Evidenzqualität

6. Fazit

Für Frauen der **Altersgruppe 40-49 Jahre** sprechen die vorliegenden Daten für einen geringen relativen Nutzen (15%) bezüglich des Endpunktes „Brustkrebssterblichkeit“:

Der absolute Nutzen bezüglich des Endpunktes Brustkrebssterblichkeit liegt bei etwa einer von ca. 2.500 Frauen, die dazu über einen Zeitraum von 10 Jahren regelmäßig zum Screening eingeladen werden müssen. Dem stehen falsch-positive Befunde kumulativ über 10 Jahre bei 20-50% aller Eingeladenen gegenüber, dazu eine noch ungewisse Anzahl von Überdiagnosen sowie möglicherweise eine dem Nutzen vergleichbare Anzahl von zusätzlich induzierten Mammakarzinomen.

Hinsichtlich der möglichen negativen Effekte besteht aufgrund spärlicher und schlecht übertragbarer Datenlage eine große Unsicherheit der Aussagen.

Dementsprechend gilt diese Unsicherheit auch für die Beurteilung des so genannten Netto-Nutzens.

Die Entscheidung für oder gegen ein Mammographie-Screening in dieser Altersgruppe wird aufgrund der dargestellten Studienlage stark von wertenden Abwägungen bestimmt werden [28;30;36].

Für Frauen der **Altersgruppe 70-74 Jahre** sprechen die vorliegenden Daten eher für einen Schaden als für einen Nutzen bezüglich des Endpunktes „Brustkrebssterblichkeit“:

Die Datenlage zu dieser Fragestellung ist allerdings sowohl hinsichtlich des potentiellen Nutzens als auch des potentiellen Schadens sehr spärlich, so dass alle Entscheidungen mit einer großen Unsicherheit behaftet sind. Den vorliegenden Daten kann kein hinreichender Beleg für einen Nutzen entnommen werden, so dass ein Mammographie-Screening in dieser Altersgruppe nicht gerechtfertigt scheint.

7. Literatur

1. Smart CR, Hendrick RE, Rutledge JH, III, Smith RA. Benefit of mammography screening in women ages 40 to 49 years. Current evidence from randomized controlled trials. *Cancer* 1995;75:1619-26.
2. Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. Efficacy of screening mammography. A meta-analysis. *JAMA* 1995;273:149-54.
3. Glasziou PP, Woodward AJ, Mahon CM. Mammographic screening trials for women aged under 50. A quality assessment and meta-analysis. *Med J Aust* 1995;162:625-9.
4. Swedish Cancer Society, Swedish National Board of Health and Welfare. Breast-cancer screening with mammography in women aged 40-49 years. *Int J Cancer* 1996;68:693-9.
5. Kerlikowske K. Efficacy of screening mammography among women aged 40 to 49 years and 50 to 69 years: comparison of relative and absolute benefit. *J Natl Cancer Inst Monogr* 1997;79-86.
6. Glasziou P, Irwig L. The quality and interpretation of mammographic screening trials for women ages 40-49. *J Natl Cancer Inst Monogr* 1997;73-7.
7. Hendrick RE, Smith RA, Rutledge JH, III, Smart CR. Benefit of screening mammography in women aged 40-49: a new meta-analysis of randomized controlled trials. *J Natl Cancer Inst Monogr* 1997;87-92.
8. Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable?. *Lancet* 2000;355:129-34.
9. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database of Systematic Reviews* 2006.
10. Ringash J. Preventive health care. 2001 update: screening mammography among women aged 40-49 years at average risk of breast cancer. *CMAJ* 2001;164:469-76.
11. Berrington G, Reeves G. Mammographic screening before age 50 years in the UK: comparison of the radiation risks with the mortality benefits. *Br J Cancer* 2005;93:590-6.
12. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. *Lancet* 2006;368:2053-60.
13. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. *Ann Intern Med* 2002;137:305-12.
14. Armstrong K, Moye E, Williams S, Berlin JA, Reynolds EE. Screening mammography in women 40 to 49 years of age: a systematic review for the American College of Physicians. *Ann Intern Med* 2007;146:516-26.
15. Qaseem A, Snow V, Sherif K, Aronson M, Weiss KB, Owens DK. Screening mammography for women 40 to 49 years of age: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 2007;146:511-5.
16. Fletcher SW, Elmore JG. False-positive mammograms - can the USA learn from Europe?. *Lancet* 2005;365:7-8.
17. de Koning HJ. Mammographic screening: evidence from randomised controlled trials. *Ann Oncol* 2003;14:1185-9.
18. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmo mammographic screening trial: follow-up study. *BMJ* 2006;332:689-92.
19. Nystrom L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Ryden S et al. Breast cancer screening with mammography: overview of Swedish randomised trials. *Lancet* 1993;341:973-8.

20. Larsson LG, Andersson I, Bjurstam N, Fagerberg G, Frisell J, Tabar L et al. Updated overview of the Swedish Randomized Trials on Breast Cancer Screening with Mammography: age group 40-49 at randomization. *J Natl Cancer Inst Monogr* 1997;57-61.
21. Cox B. Variation in the effectiveness of breast screening by year of follow-up. *J Natl Cancer Inst Monogr* 1997;69-72.
22. Berry DA. Benefits and risks of screening mammography for women in their forties: a statistical appraisal. *J Natl Cancer Inst* 1998;90:1431-9.
23. Law J, Faulkner K. Cancers detected and induced, and associated risk and benefit, in a breast screening programme. *Br J Radiol* 2001;74:1121-7.
24. Ernster VL, Barclay J, Kerlikowske K, Grady D, Henderson C. Incidence of and treatment for ductal carcinoma in situ of the breast. *JAMA* 1996;275:913-8.
25. Ernster VL, Ballard-Barbash R, Barlow WE, Zheng Y, Weaver DL, Cutter G et al. Detection of ductal carcinoma in situ in women undergoing screening mammography. *J Natl Cancer Inst* 2002;94:1546-54.
26. Moss S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis in randomised controlled trials of breast cancer screening. *Breast Cancer Res* 2005;7:230-4.
27. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. *Ann Intern Med* 2007;146:502-10.
28. Dixon JM. Screening for breast cancer. *BMJ* 2006;332:499-500.
29. Moss S. A trial to study the effect on breast cancer mortality of annual mammographic screening in women starting at age 40. Trial Steering Group. *J Med Screen* 1999;6:144-8.
30. Elmore J, Choe J. Breast cancer screening for women in their 40s: moving from controversy about data to helping individual women. *Ann Intern Med* 2007;146:529-31.
31. Hofvind S, Thoresen S, Tretli S. The cumulative risk of a false-positive recall in the Norwegian Breast Cancer Screening Program. *Cancer* 2004;101:1501-7.
32. Tabar L, Yen MF, Vitak B, Chen HH, Smith RA, Duffy SW. Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. *Lancet* 2003;361:1405-10.
33. Zahl PH, Strand BH, Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study. *BMJ* 2004;328:921-4.
34. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis. 4th ed. Luxembourg: Office for Official Publications of the European Communities; 2006.
35. Chen HH, Thurfjell E, Duffy SW, Tabar L. Evaluation by Markov chain models of a non-randomised breast cancer screening program in women aged under 50 years in Sweden. *J Epidemiol Community Health* 1998;52:329-35.
36. Djulbegovic B, Lyman GH. Screening mammography at 40-49 years: regret or no regret?. *Lancet* 2006;368:2035-7.
37. Mandelblatt J, Saha S, Teutsch S, Hoerger T, Siu AL, Atkins D et al. The cost-effectiveness of screening mammography beyond age 65 years: a systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2003;139:835-42.
38. Kerlikowske K, Salzman P, Phillips KA, Cauley JA, Cummings SR. Continuing screening mammography in women aged 70 to 79 years: impact on life expectancy and cost-effectiveness. *JAMA* 1999;282:2156-63.
39. Croyle RT. Psychosocial effects of screening for disease prevention and detection. Oxford: Oxford Univ. Pr.; 1995.
40. Landelijke evaluatie Team voor bevolkingsonderzoek naar Borstkranker (LETB). Landelijke evaluatie van bevolkingsonderzoek naar Borstkanker in nederland. Rotterdam: 2002.

41. Paci E, Ponti A, Zappa M, Patriarca S, Falini P, Delmastro G et al. Early diagnosis, not differential treatment, explains better survival in service screening. *Eur J Cancer* 2005;41:2728-34.
42. Olsen O, Gotzsche PC. Screening for breast cancer with mammography. *Cochrane Database of Systematic Reviews* 2001.
43. Giersiepen K, Hense HW, Klug SJ, Antes G, Zeeb H. Entwicklung, Durchführung und Evaluation von Programmen zur Krebsfrüherkennung Ein Positionspapier. *Z Arztl Fortbild Qualitatssich* 2007;101:43-9.
44. Law J, Faulkner K, Young KC. Risk factors for induction of breast cancer by x-rays and their implications for breast screening. *Br J Radiol* 2006;80:261-6.
45. Law J, Faulkner K. Radiation benefit and risk at the assessment stage of the UK Breast Screening Programme. *Br J Radiol* 2006;79:479-82.
46. Boer R, Warmerdam P, de Koning H, van Oortmarssen G. Extra incidence caused by mammographic screening. *Lancet* 1994;343:979.
47. Jonsson H, Johansson R, Lenner P. Increased incidence of invasive breast cancer after the introduction of service screening with mammography in Sweden. *Int J Cancer* 2005;117:842-7.
48. Paci E, Miccinesi G, Puliti D, Baldazzi P, De Lisi V, Falcini F et al. Estimate of overdiagnosis of breast cancer due to mammography after adjustment for lead time. A service screening study in Italy. *Breast Cancer Res* 2006;8:R68.
49. Kunz R, Lelgemann M, Guyatt G, Antes G, Falck-Ytter Y, Schünemann H. Von der Evidenz zur Empfehlung. In Kunz R, Ollenschläger G, Raspe H, Jonitz G, Donner-Banzhoff N, eds. *Lehrbuch Evidenz-basierte Medizin*. Köln: Deutscher Ärzteverlag; 2007. pp 231-47.
50. Guyatt G, Gutterman D, Baumann MH, Addrizzo-Harris D, Hylek EM, Phillips B et al. Grading strength of recommendations and quality of evidence in clinical guidelines: report from an american college of chest physicians task force. *Chest* 2006;129:174-81.
51. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;137:347-60.
52. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002;359:909-19.
53. Gonzalez EC, Ferrante JM, Van Durme DJ, Pal N, Roetzheim RG. Comorbid illness and the early detection of cancer. *South Med J* 2001;94:913-20.
54. Elmore JG, Barton MB, Mocerri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. *N Engl J Med* 1998;338:1089-96.
55. Haynes RB, McKibbon KA, Wilczynski NL, Walter SD, Werre SR. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005;330:1179.
56. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S et al. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328:1490.
57. Shapiro S, Venet W, Strax P, Venet L. Current results of the breast cancer screening randomized trial: the health insurance plan (HIP) of greater New York study. In: Day NE, Miller AB, eds. *Screening for Breast Cancer*. Toronto: Hans Huber; 1988. pp 3-15.
58. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. *CMAJ* 1992;147:1459-76.
59. Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50-59 years. *J Natl Cancer Inst* 2000;92:1490-9
60. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. *CMAJ* 1992;147:1477-88

61. Alexander FE, Anderson TJ, Brown HK, Forrest AP, Hepburn W, Kirkpatrick AE et al. 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *Lancet* 1999;353:1903-8.
62. Bjurstam N, Bjorneld L, Duffy SW, Smith TC, Cahlin E, Erikson O et al. The Gothenburg Breast Cancer Screening Trial: preliminary results on breast cancer mortality for women aged 39-49. *J Natl Cancer Inst Monogr* 1997;53-5.
63. Frisell J, Lidbrink E, Hellstrom L, Rutqvist LE. Followup after 11 years - update of mortality results in the Stockholm mammographic screening trial. *Breast Cancer Res Treat* 1997;45:263-70.
64. Andersson I, Janzon L. Reduced breast cancer mortality in women under age 50: updated results from the Malmo Mammographic Screening Program. *J Natl Cancer Inst Monogr* 1997;63-7.
65. Tabar L, Fagerberg G, Chen HH, Duffy SW, Smart CR, Gad A et al. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. *Cancer* 1995;75:2507-17.
66. Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grontoft O et al. Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet* 1985;1:829-32.
67. Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F et al. Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. *BMJ* 1988;297:943-8.
68. Fracheboud J, Groenewoud JH, Boer R, Draisma G, de Buijn AE, Verbeek AL et al. Seventy-five years is an appropriate upper age limit for population-based mammography screening. *Int J Cancer* 2006;118:2020-5.

A Mammographie-Screening
bei Frauen 40-49 Jahre und >70 Jahre
Nutzen/Schaden-Bilanz

B **Doppelbefundung von Screening-Mammographien unter Berücksichtigung des Entscheidungsverfahrens bei diskrepanten Befunden und des Einsatzes von Computer-Assistierten-Diagnose (CAD)-Systemen**

C Stellenwert der Mammasonographie im Rahmen der Brustkrebsfrüherkennung

D Indikationsstellung für die stereotaktische Vakuumbiopsie im Vergleich zur stereotaktischen Stanzbiopsie

E Nutzen der Exzision des Stichkanals nach minimal- invasiven Biopsie der Brust zur Vermeidung von Tumorzellverschleppung

F Her-2/neu-Bestimmung:
Vergleich der Methoden Immunhistochemie,
Fluoreszenz-in-situ-Hybridisierung und
Chromogene-in-situ-Hybridisierung

Inhalt

1. Hintergrund	39
2. Fragestellung	41
3. Methoden	43
3.1 Kriterien für den Einschluss und Ausschluss von Studien	43
3.1.1 Zielgruppe	43
3.1.2 Diagnostische Verfahren und Vergleichsverfahren	43
3.1.3 Studientypen	43
3.1.4 Erforderliche Angaben:	44
3.1.5 Einschluss- und Ausschlusskriterien	44
3.2. Literaturrecherche	44
3.2.1 Datenbanken und Suchbegriffe	44
3.3 Zielgrößen und deren Definition	45
3.4. Bewertung der gefundenen Literatur	45
4. Ergebnisse	46
4.1. Doppelbefundung von Screening-Mammographien unter Berücksichtigung der Entscheidungsverfahren	46
4.1.1 Ergebnisse der Literaturrecherche	46
4.1.2 Aufbereitete Evidenz	46
4.1.3 Evidenz von Einzelstudien	47
4.1.4 Diskussion	48
4.2 Doppelbefundung mit CAD-Systemen	49
4.2.1 Ergebnisse der Literaturrecherche:	49
4.2.2 Aufbereitete Evidenz	49
4.2.3 Evidenz von Einzelstudien	49
4.2.4 Diskussion	51
5. Evidenztabellen	52
5.1 Evidenz zu Doppelbefundung von Screening-Mammographien	52
5.1.1 Aggregierte Evidenz	52
5.1.2 Prospektive Kohortenstudien: Studiencharakteristika	53
5.1.3 Retrospektive Kohortenstudien: Studiencharakteristika	56
5.1.4 Prospektive Kohortenstudien - Ergebnisse	57
5.1.5 Retrospektive Kohortenstudien- Ergebnisse	62
5.2 Evidenz zum Einsatz von CAD-Systemen beim Mammographie-Screening	65
5.2.1 Kohortenstudien: Studiencharakteristika	65
5.2.2 Kohortenstudien: Ergebnisse	69
5.3 Ausgeschlossene Studien	72
5.3.1 Ausgeschlossene Studien zur Doppelbefundung	72
5.3.2 Ausgeschlossene Studien zur Befundung mit CAD-Systemen	72
6. Kosten-Nutzen-Analysen	74
7. Literatur	77

Tabellenverzeichnis

Tabelle 1	Aggregierte Evidenz	52
Tabelle 2a	Prospektive Kohortenstudien: Studiencharakteristika	53
Tabelle 2b	Retrospektive Kohortenstudien: Studiencharakteristika	55
Tabelle 3a	Prospektive Kohortenstudien - Ergebnisse	57
Tabelle 3b	Retrospektive Kohortenstudien - Ergebnisse	61
Tabelle 4	Kohortenstudien: Ergebnisse zu CAD	69
Tabelle 5	Ausgeschlossene Studien zur Doppelbefundung	72
Tabelle 6	Ausgeschlossene Studien zur Befundung mit CAD-Systemen	72
Tabelle 7	Kosten-Nutzen-Analysen	74

1. Hintergrund

Im populationsbezogenen Mammographie-Screening werden nicht alle zum Zeitpunkt der Screening-Untersuchung vorhandenen Karzinome erkannt. Bei Frauen, die im weiteren Verlauf Brustkrebs entwickeln, kann die Befundveränderung bei retrospektiver Betrachtung der als unauffällig befundeten Screening-Mammographien in 3% bis 25% der Fälle diagnostiziert werden [1,2]. Bei nicht erkannten Karzinomen handelt es sich überwiegend um Herdbefunde und Architekturstörungen und zu etwa einem Drittel um suspekta Mikrokalzifikationen [2]. Durch die Doppelbefundung von Mammographien soll vor allem die Anzahl falsch negativer mammographischer Befunde reduziert werden.

Die Doppelbefundung von Screening-Mammographien wurde bereits in der Stufe-3-Leitlinie zur Brustkrebsfrüherkennung in Deutschland 2003 empfohlen [3]. Diese Empfehlung sollte nun einer systematischen Evidenzüberprüfung unterzogen werden.

Verfahren der Doppelbefundung

Eine Doppelbefundung von Screening-Mammographien kann in unterschiedlichen Protokollen ausgeführt werden. Zum einen gibt es die unabhängige Doppelbefundung, d.h. die beiden Befunder kennen den jeweils anderen Befund nicht, zum anderen die nicht-unabhängige Doppelbefundung, d.h. der zweite Befunder kennt den Befund des ersten.

Hinsichtlich des weiteren Vorgehens bei diskrepanten Befunden unterscheidet man:

- a.) Unilaterales Verfahren: Wiedereinbestellung bei jedem suspektem Befund (egal ob von Befunder 1 oder Befunder 2 allein oder von beiden als suspekt bewertet).
- b.) Konsensus-Verfahren: Wiedereinbestellung nur nach Einigung der beiden Befunder über diskrepante Befunde.
- c.) ‚Arbitration‘: Wiedereinbestellung nach Entscheidung diskrepanter Befunde durch einen dritten Befunder.

Die Entscheidungsverfahren haben nicht nur Einfluss auf die Sensitivität sondern auch auf die Spezifität und die Wiedereinbestellungsrate und damit auf den positiven prädiktiven Wert des Screenings [4]. Die Doppelbefundung muss deshalb immer mit dem sich anschließenden Entscheidungsverfahren bewertet werden.

Für das deutsche Mammographie-Screening gilt die unabhängige Doppelbefundung mit ‚Arbitration‘ durch den programmverantwortlichen Arzt unter Einbeziehung der beiden Vorbefunder [5].

Einflussfaktoren auf die diagnostische Wertigkeit der Befundung von Screening-Mammographien

Die Befunder von populationsbezogenen Screening-Mammographien sind mit verschiedenen Faktoren konfrontiert, die die Validität und Reliabilität des Verfahrens deutlich beeinflussen können. Diese Faktoren sollten bei der Bewertung der Doppelbefundung bekannt sein, um ihren Einfluss prüfen zu können.

1. Die Prävalenz der vorhandenen Mammakarzinome in der gescreenten Gruppe

Überwiegend sind mammographische Normalbefunde zu interpretieren, da die Mammographien an asymptomatischen Frauen vorgenommen werden. Die Prävalenz von Mammakarzinomen bei den gescreenten Frauen liegt ca. zwischen 0,5 und 0,75%. In der ersten Screening-Runde (initiales oder prävalentes Screening) werden ca. 6-9 Karzinome pro 1.000 Frauen entdeckt, in den darauf folgenden (inzidenten) Screening-Runden ca. 4-7 pro 1.000 [6].

Auch das Screening-Intervall bestimmt die Prävalenz an Karzinomen. Das Intervall beträgt für die Screening-Programme in der Regel zwei Jahre, abweichend davon in den USA ein Jahr und in Großbritannien drei Jahre. Je kürzer das Screening-Intervall ist, desto geringer ist die Rate an neu entdeckten Karzinomen bei inzidentem Screening [6].

2. Die Anzahl der angefertigten Aufnahmen

In den Screening-Programmen werden entweder eine Aufnahme pro Brust im cranio-caudalen Strahlengang oder zwei Aufnahmen, dann zusätzlich im mediolateralen Strahlengang, angefertigt [6]. Für Großbritannien wurde aufgrund einer randomisierten Studie und den Ergebnissen des bis dahin stattgefundenen Mammographie-Screenings mit der Mammographie in einer Ebene Mitte

der 90er Jahre die Mammographie in zwei Ebenen verbindlich für das prävalente Screening eingeführt [7]. Die Mammographie in zwei Ebenen erhöht die Sensitivität für die Entdeckung von Brustkrebs und führt zu einem Rückgang unklarer Befunde. Durch eine geringere Wiedereinbestellungsrate wird der positiv prädiktive Wert der Mammographie verbessert. Besonders die Sensitivität für kleine Karzinome (<15mm) lässt sich signifikant verbessern [7-9]. Im deutschen Screening ist die Mammographie in zwei Ebenen grundsätzlich obligat [10].

3. Die technische Qualität der Mammographie

Film-Folien-Mammographien, durchgeführt mit dem aktuellen technischen Standard entsprechenden Geräten, und digitale Mammographie-Aufnahmen sind hinsichtlich der Beurteilbarkeit als gleichwertig zu betrachten [11]. Mängel bei der technischen Qualität der Mammographie führen zu einer schlechteren Beurteilbarkeit und aufgrund dessen eventuell zu einer häufigeren Wiedereinbestellung.

4. Die Ausbildung und Erfahrung der Befunder

Für die Beurteilung von Screening-Mammographien ist eine spezielle Ausbildung erforderlich [5,12]. Dass erfahrener Befunder weniger Karzinome übersehen, ist eine allgemeine Annahme. Die Beurteilung der gleichen Mammographien kann jedoch von Befunder zu Befunder erheblich variieren, auch wenn diese sehr erfahren sind [13].

Computer-Assistierte-Diagnose (CAD)-Systeme

Computer-Assistierte-Diagnose-Systeme markieren veränderte Regionen an digitalen Mammographien und sollen dadurch zu einer verbesserten Karzinomdetektion beitragen. CAD-Systeme erkennen suspekte Mikrokalzifikationen mit einer höheren Treffsicherheit als Herdbefunde oder Architekturstörungen, entsprechend der ‚menschlichen‘ Befundung.

Die CAD-Markierungen weisen insgesamt eine hohe Sensitivität bei sehr geringer Spezifität auf [14].

Seit einigen Jahren wird der Einsatz dieser Systeme als Ersatz für einen zweiten Befunder auch für das Mammographie-Screening diskutiert. Diese Diskussion entstand z.B. in Großbritannien aufgrund eines Mangels an gut ausgebildeten Radiologen [15]. In Großbritannien werden zum Teil Röntgenassistenten – nach entsprechender Schulung – für die Befundung von Screening-Mammographien eingesetzt [16].

Vor diesem Hintergrund sollten der Nutzen beziehungsweise der Schaden der Doppelbefundung von Mammographien unter Berücksichtigung des weiteren Entscheidungsverfahrens im populationsbezogenen Mammographie-Screening analysiert werden. Die oben beschriebenen Faktoren (1-4) sollten in Bezug auf ihren Einfluss auf die Doppelbefundung überprüft werden.

In einem zweiten Schritt sollte untersucht werden, ob Computer-Assistierte-Diagnose-Systeme bei der Doppelbefundung den Zweitbefunder ersetzen können.

Literatur

1. Boyer B, Hauret L, Bellaiche R, Gräf C, Bourcier B, Fichet G. Cancers rétrospectivement visibles: revue de la littérature. J Radiol 2004;85;2071-8
2. Ikeda DM, Anderson I, Wattsgard C, Janzon L, Linell F. Interval Carcinomas in the Malmo Screening Trial: radiographic appearance and Prognostic considerations. AJR Am J Roentgenol 1992 Aug;159(2):287-94
3. Schulz KD, Albert US (Hrsg.). Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland. München, Wien, New York: Zuckschwerdt Verlag; 2003
4. Dinnes J, Moss S, Melia J, Blanks R, Song F, Kleijnen J. Effectiveness and cost-effectiveness of double reading of mammograms in breast cancer screening: findings of a systematic review. Breast 2001;10(6):455-63
5. Kassenärztliche Bundesvereinigung. Einführung eines bundesweiten Mammographie-Screening Programms (Programmrichtlinien). Beilage zum Deutschen Ärzteblatt Heft 4, 23.1.2004, Ausgabe A
6. International Agency for Research on Cancer (IARC): Screening for breast cancer, IARC Press, 2002

7. Engel J, Hölzel D, Schubert-Fritschle G. Epidemiologie. IN: Sauer H. (Hrsg.). Manual Mammakarzinome: Empfehlungen zur Diagnostik Therapie und Nachsorge. 10. Auflage. München, Wien, New York: Zuckschwerdt Verlag 2005
8. Wald NJ, Murphy P Major P, Parkes, Chris Frost C. UKCCCR multicentre randomised controlled trial of one and two view mammography in breast cancer screening. *BMJ* 1995 Nov 4;311(7014):1189-93
9. Given-Wilson RM, Blanks RG. Incident screening cancers detected with a second mammographic view: pathological and radiological features. *Clin Radiol* 1999 Nov;54(11):724-35
10. Blanks GR, Moss SM, Wallis MG. Use of two view mammography compared with one view in the detection of small invasive cancers: further results from the National Health Service breast screening programme. *J Med Screen* 1997;4(2):98-101
11. Skaane P, Skjennald A, Young K, Egge E, Jepsen I, Sager EM, Scheel B, Søvik E, Ertzaas AK, Hofvind S, Abdelnoor M. Follow-up and final results of the Oslo I Study comparing screen-film mammography and full-field digital mammography with soft-copy reading. *Acta Radiol.* 2005 Nov;46(7):679-89
12. Liston J, Wilson R. Quality Assurance Guidelines for Breast Cancer Screening Radiology. NHSBSP Publication No 59,.NHS Cancer Screening Programms,Sheffield.1/2005
13. Hukkinen K, Kivisaari L, Vehmas T. Impact of the number of readers on mammography interpretation. *Acta radiol*,2006; 47/7; 655-9
14. Balleyguier C, Boyer B, Athanasiou A, Vanel D, Sigal R. [Understanding CAD (computer-aided diagnosis) in mammography]. *J Radiol* 2005;86(1):29-35
15. Astley SM, Gilbert FJ. Computer-aided detection in mammography. *Clin Radiol* 2004;59(5):390-9
16. Pauli R, Hammond S, Cooke J, Ansell J. Comparison of radiographer/radiologist double film reading with single reading in breast cancer screening. *J Med Screen* 1996;3(1):18-22

2. Fragestellung

Folgende Schlüsselfragen wurden von der Methodikerin in Abstimmung mit dem Auftraggeber generiert:

1. Welchen diagnostischen Nutzen bzw. Schaden hat die Doppelbefundung von Screening-Mammographien unter Berücksichtigung der verschiedenen Entscheidungsverfahren bei nicht übereinstimmenden Befunden?
2. Welchen diagnostischen Nutzen bzw. Schaden hat der Einsatz von Computer-Assistierten-Diagnose-Systemen zur Doppelbefundung im Rahmen des Mammographie-Screenings?

Unter Nutzen wird dabei jeweils eine verbesserte Befundung von Screening-Mammographien verstanden, die in der Konsequenz zu einem Rückgang an Brustkrebsmortalität führt (siehe auch 3.1.3 Studientypen und 3.3 Zielgrößen). Unter Schaden werden zusätzliche falsch-positive Befunde verstanden, die zu einer höheren Rate an unnötigen Wiedereinbestellungen von Frauen sowie zu einer erhöhten Rate an unnötigen Biopsien führen. Unter Schaden wird darüber hinaus eine ‚Überdiagnostik‘ an Mammakarzinomen verstanden. Das bedeutet, dass zusätzliche Mammakarzinome diagnostiziert werden, die Brustkrebsmortalität aber nicht gesenkt wird [17]. Des Weiteren wurde mit dem Auftraggeber vereinbart, Kosten-Nutzen-Analysen gesondert darzustellen. Diese wurden methodisch nicht bewertet und sind unter 6. aufgeführt.

Literatur

17. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ*. 2006 Mar 25;332(7543):689-92

3. Methoden

3.1 Kriterien für den Einschluss und Ausschluss von Studien

Die folgenden Kriterien wurden prospektiv von der Methodikerin in Absprache mit dem Auftraggeber festgelegt.

3.1.1 Zielgruppe

Zielgruppe waren Frauen, die an einem organisierten populationsbezogenen Mammographie-Screening teilnehmen.

3.1.2 Diagnostische Verfahren und Vergleichsverfahren

Für die erste Frage waren folgende diagnostische Verfahren Gegenstand der vergleichenden Untersuchung: die Doppelbefundung von Screening-Mammographien mit anschließendem unilateralem Verfahren, Konsensus-Verfahren oder ‚Arbitration‘.

In Bezug auf Frage 2 war die Doppelbefundung von Screening-Mammographien mit einem CAD-System als Zweitbefunder das untersuchte diagnostische Verfahren. Das Vergleichsverfahren war die Doppelbefundung durch zwei Radiologen.

Der Referenzstandard für die Befundungsqualität war die histologische Sicherung bzw. eine Nachbeobachtung bis zum nächsten Screening. Nur die Kenntnis im Screening-Intervall aufgetretener maligner Befunde bzw. die beim nächsten Screening retrospektiv als sichtbar klassifizierten Befunde erlaubt die Berechnung falsch negativer Befunde. Dabei sind echte Intervallkarzinome von falsch-negativen Befunden zu unterscheiden.

3.1.3 Studientypen

Zur Beantwortung der Fragestellung Nutzen und Schaden der Doppelbefundung bzw. verschiedener Doppelbefundungsstrategien sind grundsätzlich randomisierte kontrollierte Studien mit dem Endpunkt Brustkrebsmortalität geeignet. Nur in solchen Studien kann gezeigt werden, inwieweit diese Verfahren zu einer Überdiagnostik führen. Orientierende Vorrecherchen ergaben, dass zum Recherchezeitpunkt keine randomisierten kontrollierten Studien zu der fokussierten Fragestellung vorlagen. Deshalb wurde im nächsten Schritt nach prospektiven Kohortenstudien gesucht. Auch retrospektive Kohortenstudien wurden in die Auswertung miteinbezogen, wenn sie den Einschlusskriterien entsprachen. Fall-Kontroll-Studien wurden für diese Fragestellung nicht in die Evidenzbewertung eingeschlossen. Die Ergebnisse dieser Studien sind durch die um ein Vielfaches höhere Prävalenz von Mammakarzinomen (die Rate Fälle zu Kontrollen liegt in der Regel zwischen 1:1 bis 1:5) nicht direkt auf die reale Screening-Situation übertragbar, bei der eine Karzinomprävalenz von maximal 0,9% vorliegt. Fall-Kontroll-Studien können in Bezug auf die Befundung von Screening-Mammographien nur hypothesengenerierende Funktion haben.

Aufbereitete Evidenz – systematische Reviews – wurde vorrangig eingeschlossen, wenn die darin enthaltenen Studien den Einschlusskriterien entsprachen.

Kosten-Nutzenanalysen:

Die Ergebnisse von Kosten-Effektivitätsanalysen werden unter 6. gesondert aufgeführt und das zugrunde liegende Verfahren kurz beschrieben. Eine Bewertung der Validität der Angaben erfolgte nicht. Die Ergebnisse werden deshalb nicht dargestellt.

3.1.4 Erforderliche Angaben:

Die zum Studieneinschluss erforderlichen Angaben wurden von der Methodikerin im Hinblick auf die Auswertbarkeit und Vergleichbarkeit der Studien vorab festgelegt:

- Screening-Situation (prävalent, inzident)
- Altersgruppe
- Aufnahmezahl (Mammographie in einer oder in zwei Ebenen)
- Angabe der Wiedereinbestellungsrate und/oder der Karzinomentdeckungsrate bei Einzel- oder Doppelbefundung
- Angabe des Entscheidungsverfahrens nach Doppelbefundung: Unilateral, Consensus, Arbitration

3.1.5 Einschluss- und Ausschlusskriterien

Im Hinblick auf die oben aufgeführten Einflussfaktoren und die festgelegten Endpunkte wurden folgende **Einschlusskriterien** festgelegt:

1. Studie/Review beantwortet die Fragestellung
2. adäquater Studientyp
3. adäquate Studienpopulation
4. erforderliche Angaben sind genannt

Alle Einschlusskriterien mussten erfüllt sein.

Es galten folgende **Ausschlusskriterien**:

1. Doppelpublikation
2. Methodisch: Fall-Kontroll-Studien, Fallberichte und Expertenmeinungen

Kein Ausschlusskriterium durfte vorliegen.

3.2. Literaturrecherche

Die Literaturrecherche erfolgte nach Absprache mit dem Auftraggeber ohne zeitliche Beschränkung. Rechercheschlussdatum war der 16.2.2007. Eine Nachrecherche am 10.5.2007 erbrachte eine neue Publikation für die Einzelbefundung mit CAD-Systemen (Fenton et al., 2007).

Die identifizierten Abstracts wurden von der Methodikerin auf inhaltliche Relevanz geprüft, nicht relevante Abstracts wurden ausgeschlossen. Inhaltlich als relevant bewertete Abstracts wurden als Volltexte beurteilt. Der Ausschluss von Volltexten ist jeweils begründet.

3.2.1 Datenbanken und Suchbegriffe

In folgenden Datenbanken wurde nach Vorgabe des Auftraggebers im ÄZQ eine systematische Literaturrecherche durchgeführt:

1. PubMed (Internetportal der National Library of Medicine)
(<http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed>)
2. Database of Abstracts of Reviews of Effects (DARE) des Centre for Reviews and Dissemination (CRD) (<http://www.york.ac.uk/inst/crd/crddatabases.htm#DARE>)
3. Cochrane-Datenbanken 'Cochrane Reviews' und 'Clinical Trials'
(<http://www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/HOME>).

Folgende Suchbegriffe wurden verwendet:

a) Für PubMed:

- 1) (double OR dual OR second) AND reading* AND (breast OR mammo*)
jeweils als Textwort, d.h. nicht nur auf Titel oder abstract bezogen.

Eine Suche mit Mesh-Terms (mass screening, breast neoplasms) ergab keine zusätzlichen Zitate.

- 2) (computer-aided detection OR computer aided detection OR CAD) AND Screening AND breast AND (detection or recall OR false negative)
jeweils als Textwort.
Die Suchworte schließen hier auch ergebnisorientierte Schlagworte mit ein, da zu CAD viele experimentelle Studien existieren, die nicht Gegenstand der Fragestellung waren.

b) Für DARE und Cochrane:

- 1) ('double reading' AND 'mammography screening' OR 'breast Cancer screening')
- 2) ('computer aided diagnosis' AND 'mammography screening' OR 'breast cancer screening')

3.3 Zielgrößen und deren Definition

Verlässliche Aussagen für die Effektivität der Doppelbefundung können durch eine erniedrigte Mortalitätsrate an Mammakarzinomen gezeigt werden. Da die orientierende Vorrecherche zeigte, dass zum Recherchezeitpunkt zu diesem Endpunkt keine Studien vorlagen, wurden folgende Testeigenschaften für diagnostische Tests [18] als Surrogatparameter für die Bewertung der Doppelbefundung festgelegt (Definition siehe S.91):

- ⇒ Positiver prädiktiver Wert (PPV)
- ⇒ Negativer prädiktiver Wert (NPV)
- ⇒ Sensitivität
- ⇒ Spezifität

Für die Berechnung dieser Testeigenschaften der mammographischen Befundung wurde die Anzahl diagnostizierter Mammakarzinome aufgrund suspekt befundeter Mammographien herangezogen. Der Anteil als suspekt befundeter Mammographien wurde mit der Rate wiedereinbestellter Frauen gleichgesetzt.

3.4. Bewertung der gefundenen Literatur

Die identifizierten Studien wurden nach Vorgabe des Auftraggebers mit der Oxford-Klassifikation für Diagnosestudien bewertet (siehe Anhang 1).

Literatur

18. Gordis L. Epidemiologie. Marburg: Kilian-Verlag 2001

4. Ergebnisse

4.1. Doppelbefundung von Screening-Mammographien unter Berücksichtigung der Entscheidungsverfahren

4.1.1 Ergebnisse der Literaturrecherche

Die Recherche in DARE und Cochrane ergab insgesamt neun Zitate. Fünf der Abstracts erfüllten nicht die inhaltlichen Einschlusskriterien, vier waren Dubletten zu PubMed.

Die PubMed-Suche ergab 195 Einzel-Zitate. Von den 195 Zitaten wurden aufgrund der Abstracts 19 Studien zur Doppelbefundung identifiziert. Zwei zusätzliche Kohortenstudien wurden aufgrund des systematischen Reviews von Dinnes et al. identifiziert (diese wurden in ‚Breast‘ publiziert, bevor die Zeitschrift für PubMed indexiert wurde).

Rechercheergebnisse für PubMed zur Doppelbefundung

PubMed bis 16.2.2007	n= 195
Nach Abstractscreening geprüfte Volltexte	n= 22
Systematischer Review	n= 1
Kohortenstudien prospektiv	n= 9 (2 nach 7/1999)
Kohortenstudien retrospektiv	n= 6 (2 nach 7/1999)
Fall-Kontroll-Studien	n= 4
Handsuche	n= 2 (durch Review)
Eingeschlossene Einzelstudien	n= 4 (nach 7/1999 publiziert)

4.1.2 Aufbereitete Evidenz

Eingeschlossen wurde der systematische Review: ‚Effectiveness and Cost-Effectiveness of Double Reading of Mammograms in Breast Cancer Screening‘ von J. Dinnes et al., 2001.

Dieser Review ist im Volltext nicht publiziert worden, der Volltext wurde aber auf Nachfrage vom CRD (Centre of Reviews and Dissemination) zur Verfügung gestellt und kann im ÄZQ eingesehen werden.

a) Studien- und Publikationsqualität

Von den Autoren wurde eine umfassende systematische Literatursuche in relevanten Datenbanken (u.a. Medline, Embase, DARE, CancerLit) bis Juli 1999 durchgeführt.

Es wurden stringente Einschlusskriterien definiert. Die Daten von sieben prospektiven und zwei retrospektiven Kohortenstudien (als Studientyp A bezeichnet) wurden für die Analyse der Doppelbefundung vorrangig herangezogen¹. Alle diese Studien wurden in nationalen oder regionalen Screening-Programmen durchgeführt.

Für vier der Studien lagen Follow-up-Daten vor, davon für drei bis zur nächsten Screening-Runde. Die Studie mit unvollständigem Follow-up (Pauli et al., 1996) wurde für die Berechnung der Sensitivität und Spezifität der Doppelbefundung nicht herangezogen.

Neben den Testeigenschaften wurde für alle Studien die Rate entdeckter Karzinome und die Rate wiederbestellter Frauen angegeben. Keine der Studien hatte als Endpunkt die Brustkrebsmortalitätsrate.

Fünf weitere Studien wurden als qualitativ schlechter eingestuft (Typ B), dies waren drei retrospektive Kohortenstudien mit unvollständigen Angaben und zwei Fall-Kontrollstudien. Die Ergebnisse dieser Studien sind im Review ebenfalls dargestellt².

Die Ergebnisse der Einzelstudien wurden nicht im Sinne einer Meta-Analyse korreliert, da die Studien als zu different eingeschätzt wurden.

¹ Die Vollpublikation wurde erst nach Fertigstellung der Evidenzbewertung der Einzelstudien, die nach prospektiven und retrospektiven Studien getrennt vorgenommen wurde, zur Verfügung gestellt. Alle Studien sind deshalb in den Evidenztabelle detailliert beschrieben. Die als Studientyp A bezeichneten Studien des systematischen Reviews sind in den Evidenztabelle mit ‚*‘ gekennzeichnet.

² Die als Studientyp B klassifizierten Studien des Reviews sind mit einem # gekennzeichnet.

Die Publikationsqualität der Vollpublikation ist hinsichtlich der Angaben aus den Einzelstudien als sehr gut zu werten.

Der Review wurde nach der Oxford-Klassifikation mit 2a bewertet.

b) Inhaltliche Ergebnisse

Die Studienpopulationen lagen zwischen $n= 15.457$ und $n= 256.003$. Es wurden Frauen in der Altersgruppe zwischen 50-75 Jahren gescreent.

Die Autoren kamen zu folgenden Ergebnissen:

Karzinomdetektion

Die Rate entdeckter Karzinome erhöhte sich in jeder Studie pro Befunder um 0,01% bis 0,2%, d.h. kein Befunder erzielte in der Einzelbefundung eine höhere Rate entdeckter Karzinome als in der Doppelbefundung, unabhängig vom anschließenden Entscheidungsverfahren.

Wiedereinbestellungsraten

Die Wiedereinbestellungsraten bei der Doppelbefundung lagen im Median bei 4,5% (2,5-7,7%).

- a) Bei unilateralem Verfahren
Durch die Doppelbefundung mit unilateralem Entscheidungsverfahren (vier Studien) blieb die Wiedereinbestellungsrate pro Befunder entweder gleich oder erhöhte sich zwischen 0,4% bis zu 2,7%.
- b) Bei Konsensus-Verfahren
Bei Konsensus-Verfahren (zwei Studien) ergab sich in einer Studie eine erniedrigte Wiedereinbestellungsrate von -0,3%, in der zweiten Studie stieg die Rate wiedereinbestellter Frauen um 0,9%.
- c) Bei Arbitration
Konsensus plus Arbitration oder nur Arbitration als Entscheidungsverfahren führte in drei Studien zu einer gleichbleibenden oder niedrigeren Rate an wiedereinbestellten Frauen von -0,4% bis zu -2,7%.

Testgüte

Der Zugewinn an Sensitivität lag zwischen 2,9% und 13,7%, im Median bei 7,8%.

Die Spezifität lag für alle Befunder über 90%.

Durch die Doppelbefundung aufgrund von Konsensus und Arbitration erhöhte sich die Spezifität um 0,3% bis zu 2,8%. Bei unilateralem Verfahren erniedrigte sich die Spezifität um bis zu 2,1%.

Der positive prädiktive Wert für die Doppelbefundung von Mammographien verbesserte sich bei Konsensus- oder Arbitrationverfahren und verschlechterte sich bei unilateraler Wiedereinbestellung.

Der negative prädiktive Wert verbesserte sich bei jeder Art der Doppelbefundung.

Testgüte in Abhängigkeit weiterer Parameter

In Bezug auf das Alter der gescreenten Frauen, die Verwendung von Mammographien in einer oder in zwei Ebenen, die Anzahl oder Erfahrung der Befunder oder die Screening-Runde, ließ sich keine eindeutige Korrelation zu dem durch Doppelbefundung gefundenen Zugewinn an entdeckten Karzinomen oder Sensitivität aufzeigen.

Biopsieraten

Die angegebenen Biopsieraten bei Doppelbefundung lagen zwischen 0,8% und 0,92%.

4.1.3 Evidenz von Einzelstudien

a) Studien- und Publikationsqualität

Zu der vorliegenden Fragestellung wurden bis Juli 1999 (Recherchezeitraum v. Dinnes et al.) keine weiteren Einzelstudien identifiziert.

Ab August 1999 wurden zwei prospektive und zwei retrospektive Kohortenstudien identifiziert, die die Einschlusskriterien erfüllten.

Ausgeschlossen wurden eine retrospektive Kohortenstudie und zwei retrospektive Fall-Kontrollstudien (siehe 5.2.1 und 5.2.2).

Alle Studien enthielten Ergebnisse aus populationsbezogenen regionalen oder nationalen Screening-Programmen. In drei der Studien wurde jeweils eine Doppelbefundungsstrategie bewertet, zweimal das unilaterale Verfahren und ein Mal das Arbitrationverfahren.

Eine niederländische Studie verglich die verschiedenen Doppelbefundungsstrategien. Aufgrund der Oxford-Klassifikation wurde für diese Studie (Dujim et al., 2004) mit einer Nachbeobachtung von mindestens einem Screening-Intervall und verblindeter Doppelbefundung ein Evidenzlevel 2b vergeben.

Für die anderen Studien – ohne Follow-up – wurde ein Evidenzlevel 3b vergeben.

Es wurde keine Studie mit dem Endpunkt Brustkrebsmortalität identifiziert.

In Bezug auf die Publikationsqualität ist anzumerken, dass die Anzahl entdeckter Karzinome für alle Studien angegeben war. Die Anzahl wiedereinbestellter Frauen war für eine Studie nicht ermittelbar.

b) Inhaltliche Ergebnisse

Die Studienpopulationen lagen zwischen $n= 28.231$ und $n= 177.631$.

Die Rate entdeckter Karzinome erhöhte sich in jeder Studie durch Doppelbefundung.

Für die beiden Studien zum unilateralen Verfahren der Doppelbefundung – eine aus dem Screening-Projekt von Singapore (Kwek et al., 2003) mit verblindeter Doppelbefundung und eine italienische regionale Studie (Ciatto et al., 2005) mit nicht verblindeter Doppelbefundung bestätigte sich die Analyse von Dinnes et al. für das unilaterale Verfahren: die Wiedereinbestellungsraten erhöhten sich um 0,44 bis 2,61%.

Bei Dujim et al., 2004, die verschiedene Entscheidungsverfahren nach Doppelbefundung untersuchten, lag die Wiedereinbestellungsrate zwischen 1,1% bei unilateralem Entscheidungsverfahren und 0,09% bei Konsensus-Verfahren mit Arbitration.

Sie gaben für das unilaterale Verfahren eine Sensitivität von 76,6% und für das Konsensus-Verfahren eine kaum erniedrigte Sensitivität von 76,3% an. Ein zusätzliches Arbitration-Verfahren bei diskrepanter Doppelbefundung senkte die Sensitivität auf 75,6%, drei Karzinome wurden nicht erkannt. Der negative prädiktive Wert blieb mit 0,99 gleich. Die Spezifität ließ sich nicht vergleichend berechnen.

4.1.4 Diskussion

a) Methodische Diskussion

Für die vorliegende Fragestellung zum Stellenwert der Doppelbefundung konnte ein systematischer Review mit einer Literaturrecherche bis 1999 identifiziert werden. Die wesentlichen Befunde dieses Reviews stützen sich auf neun Kohortenstudien. Drei dieser Studien waren prospektive Studien, für die eine Nachbeobachtung mit Identifizierung von Intervall-Karzinomen bis mindestens zum nächsten Screening vorlag. Die Evidenzrecherche erbrachte für Studien nach 1999 lediglich eine weitere prospektive Studie mit Angaben zu Sensitivität und Spezifität aufgrund eines entsprechenden Follow-ups. In dieser Studie wurden Doppelbefundungsstrategien verglichen. Die aktuell identifizierten drei weiteren retrospektiven Kohortenstudien konnten zum Vergleich der durch den Review gefundenen Ergebnisse verwendet werden. Alle Studien wurden in populationsbezogenen Screening-Programmen durchgeführt, die Ergebnisse sind daher auf solche Programme anwendbar.

b) Inhaltliche Diskussion

Eine abschließende Bewertung des Nutzens von Doppelbefundungsverfahren kann aufgrund der vorliegenden Studienlage nicht gegeben werden, da keine der Studien Angaben zu den Auswirkungen auf die Brustkrebsmortalität macht.

Für die Testeigenschaften Sensitivität und Spezifität des Mammographie-Screenings konnte gezeigt werden, dass die Doppelbefundung die Sensitivität des Mammographie-Screenings um 2,9-13,7% erhöht (im Median um 7,8%).

Für das unilaterale Verfahren resultierte eine geringere Spezifität von bis zu 2,1%. Für das Konsensus- und Arbitrationverfahren erhöhte sich die Spezifität um bis zu 2,8%.

Die nach 1999 identifizierten Studien bestätigten die bisherigen Ergebnisse.

Eine Studie aus den Niederlanden (Dujim et al., 2004) bewertete den zusätzlichen Effekt von Arbitration bei einer Konsensus-Doppelbefundung negativ, weil dadurch Karzinome übersehen wurden, was eine Verschlechterung der Sensitivität von einem Prozent zur Folge hatte. Auffallend war dabei die insgesamt niedrige Sensitivität im niederländischen Screening von 76-77% bei einer sehr geringen Wiedereinbestellungsrate und einem Anteil von nur ca. 30% an Mammographien in zwei

Ebenen. Zur Verbesserung der Sensitivität kann die Wahl des Doppelbefundungsverfahrens hier nur einen sehr begrenzten Beitrag leisten.

Dass durch das Arbitration-Verfahren möglicherweise nicht die maximal mögliche Sensitivität erzielt wird, ist jedoch grundsätzlich bei der Wahl des Doppelbefundungsverfahrens zu bedenken.

Der mögliche Schaden der Doppelbefundung ist in einer erhöhten Wiedereinbestellungsrate und erhöhten Biopsieraten mit zusätzlichen negativen Befunden zu sehen. Eine höhere Anzahl an wiedereinbestellten Frauen mit einer Verschlechterung des positiven prädiktiven Werts der Mammographie-Befundung zeigte sich vor allem für das unilaterale Verfahren. Das Arbitration-Verfahren und zum Teil auch das Konsensus-Verfahren senkten den Anteil wiedereinbestellter Frauen, der positive prädiktive Wert verbesserte sich.

Die angegebenen Biopsieraten lagen unabhängig vom Verfahren der Doppelbefundung alle unter 1%. Diese Rate erscheint insgesamt niedrig. Im Hinblick auf einen Schaden durch Überdiagnostik, also Diagnosestellung für die Senkung der Brustkrebsmortalität nicht relevanter Karzinome, kann aus den vorliegenden Studien keine Aussage getroffen werden.

4.2 Doppelbefundung mit CAD-Systemen

4.2.1 Ergebnisse der Literaturrecherche:

Die systematische Literaturrecherche ergab in DARE acht Zitate, davon zwei Doppelpublikationen zu PubMed. Die weiteren Abstracts erfüllten nicht die inhaltlichen Einschlusskriterien. In Cochrane wurde kein Zitat gefunden.

Rechercheergebnisse PubMed

PubMed bis 16.2.2007	n= 289
Nach Screenen der Abstracts geprüfte Volltexte	n= 13
Reviews	n= 3
Systematische Reviews	n= 1
Kohortenstudien prospektiv	n= 6
Kohortenstudien retrospektiv	n= 3 (1 am 10.5.bei Nachreche aufgenommen, von 4/07)

4.2.2 Aufbereitete Evidenz

Es wurde kein systematischer Review identifiziert, der die Einschlusskriterien (Evaluierung von Kohortenstudien im Rahmen des Mammographie-Screenings) erfüllte. Der systematische Review vom Irwiq et al., 2004 wurde nicht eingeschlossen, weil zur Bewertung von CAD-Systemen im Mammographie-Screening auch Fall-Kontroll-Studien herangezogen wurden. Zwei weitere Reviews, die CAD-Systeme umfassend beschrieben, wurden als narrative Reviews bewertet (Astley et Gilbert, 2004 und Balleyguier et al., 2005). Sie wurden als Hintergrundliteratur verwendet.

4.2.3 Evidenz von Einzelstudien

a) Studien- und Publikationsqualität

Zu der vorliegenden Fragestellung zum Stellenwert der CAD-Befundung als Ersatz für einen Zweitbefunder konnten lediglich zwei Kohortenstudien identifiziert werden, die die Doppelbefundung mit einem CAD-System als ‚Zweitbefunder‘ und die unabhängige ‚menschliche‘ Doppelbefundung (Khoo et al., 2005 und Gilbert et al., 2006) vergleichen.

Lediglich diese beiden Studien bieten einen validen Vergleich der beiden Verfahren.

Khoo et al., 2005, führten eine prospektive Kohortenstudie durch. Gilbert et al., 2006, verwendeten bereits doppelbefundete Mammographien und werteten diese erneut mit einem Befunder und einem CAD-System aus.

Weitere fünf prospektive und zwei retrospektive Kohortenstudien verglichen die Einzelbefundung von Screening-Mammographien ohne und mit CAD-System. Diese Studien erlauben nur einen indirekten Vergleich zu Studien mit Doppelbefundungsstrategien.

Zwei dieser Studien wiesen ein Follow up von mindestens einem Screening-Intervall auf, so dass Sensitivität und Spezifität angegeben werden konnten.

Die Befundung erfolgte zunächst immer ohne Kenntnis des CAD-Befunds. Es handelte sich bei allen Studien um konsekutive Erhebungen. Die vorliegenden Studien wurden mit einem Evidenzgrad von 3b bewertet, bis auf die beiden Studien mit durchgängig erfülltem Referenzstandard, diese wurden mit einem Evidenzgrad von 2b bewertet (siehe 5.2.2).

Die ausgeschlossenen Fall-Kontroll-Studien zeigten eine Karzinomprävalenz von knapp 17% bis 100% und können deshalb in ihren Ergebnissen nicht ohne weiteres auf das Mammographie-Screening übertragen werden (siehe 5.2.3).

Hinsichtlich einer Schulung in Bezug auf das CAD-System machten lediglich zwei Studien Angaben.

b) Inhaltliche Ergebnisse

Die Studienpopulationen betragen zwischen $n = 2.389$ und $n = 2.221.355$.

Die untersuchten Altersgruppen lagen zwischen 50-74 Jahren.

Anteil zusätzlich entdeckter Karzinome und Wiedereinbestellungsrate

Khoo et al., gaben für die Doppelbefundung mit Arbitration eine Karzinomentdeckungsrate von 0,99% an und für die Befundung mit CAD einen Wert von 0,89%, d.h. eine Verschlechterung von 0,1% an.

Die Rate wiedereinbestellter Frauen erhöhte sich von 9,8% auf 10,5%.

Gilbert et al., erreichten durch CAD-Befundung eine verbesserte Karzinomentdeckung von 0,04% (1,11%) im Vergleich zur Doppelbefundung mit Arbitration- und Konsensusverfahren.

Die Rate wiedereinbestellter Frauen stieg von 6,5% auf 8,6%.

Testgüte

Für die beiden direkten Vergleichsstudien kann nur der positive prädiktive Wert angegeben werden. Dieser verschlechterte sich bei Khoo et al. von 0,1 auf 0,085 durch den Einsatz von CAD-Systemen im Vergleich zur Doppelbefundung. Auch bei Gilbert et al. verschlechterte sich der Wert gering auf 0,13 (vorher 0,14). Sensitivität, Spezifität und negativer prädiktiver Wert können nicht angegeben werden.

Anzahl und Auswahl von CAD-Markierungen

Von neun angezeigten, richtigen CAD-Markierungen wurden bei Khoo et al. sieben durch die Befunder nicht ausgewählt. Das CAD-System zeigte insgesamt 84% der entdeckten Karzinome richtig an. Es ergaben sich pro Fall 1,59 falsch positive Markierungen durch CAD.

Testgüte bei Einzelbefundung mit CAD

Für die Studien, die Einzelbefundung ohne und mit CAD-Markierung verglichen, verschlechterte sich der positive prädiktive Wert in vier Studien und blieb in drei Studien gleich. Sensitivität und Spezifität der Befundung können für zwei dieser Studien angegeben werden: die Sensitivität verbesserte sich um 3,6% (Fenton et al., 2007) und um 4% (Ko et al., 2006). Die Spezifität verschlechterte sich bei Fenton et al. um 3%, während sie bei Ko et al. gleich blieb.

Der Anteil zusätzlicher Karzinome durch die Einzelbefundung mit CAD lag bei 0,005-0,07%.

Die Rate wiedereinbestellter Frauen war bei Verwendung von CAD-Systemen bei allen Studien erhöht, im Median um 1,6% (1,1% - 3,1%).

4.2.4 Diskussion

a) Methodische Diskussion

Für die vorliegende Fragestellung zum Nutzen von CAD-Systemen für die Doppelbefundung im Mammographie-Screening wurden lediglich zwei Studien identifiziert, die die Befundung mit CAD-Systemen direkt mit einer ‚menschlichen‘ Doppelbefundungsstrategie verglichen. Für beide Studien war kein entsprechendes Follow-up gegeben, so dass Sensitivität und Spezifität nicht angegeben werden können

Aufgrund der weiteren vorliegenden sieben Kohortenstudien kann für den Einsatz von CAD-Systemen als Zusatz zur Einzelbefundung eine Aussage getroffen werden. Diese erlaubt nur einen indirekten Vergleich zu den Doppelbefundungsstrategien.

Für den Nutzen oder Schaden von CAD-Systemen im Vergleich zur Doppelbefundung ist deshalb eine valide Aussage aufgrund von nur zwei vorliegenden direkten Vergleichsstudien mit heterogenen Ergebnissen nicht möglich.

b) Inhaltliche Diskussion

Aufgrund der derzeitigen Studienlage kann nicht geschlossen werden, dass der Einsatz von CAD-Systemen den Zweitbefunder bei der Doppelbefundung ersetzen kann. Die vorliegenden wenigen Daten aus zwei direkten Vergleichsstudien zeigen ein Mal ein besseres und ein Mal ein schlechteres Ergebnis für die Doppelbefundung. Die prospektive Studie von Khoo et al. zeigte einen Nachteil für die ‚Zweitbefundung‘ mit CAD im Vergleich zur Doppelbefundung mit Arbitrationverfahren. Bei um 0,7% höherer Wiedereinbestellungsrate wurden 0,1% weniger Karzinome bei dem Einsatz eines CAD-Systems statt eines Zweitbefunders entdeckt. Die zweite Studie von Gilbert et al. zeigte einen Vorteil in der Zahl der entdeckten Karzinome und einer Mischung aus unilateralem Verfahren und Arbitration-Verfahren bei der Doppelbefundung.

In einem indirekten Vergleich der Einzelbefundung mit CAD-Systemen und der Doppelbefundung zeigt sich für die Doppelbefundung ein etwas höherer absoluter Anteil zusätzlich entdeckter Karzinome im Vergleich zu Einzelbefundung (0,005-0,07% bei Einsatz von CAD-Systemen, 0,02-0,1% bei Doppelbefundung).

Die vorliegenden Daten - eine um 3,6% und 4% verbesserte Sensitivität bei Einsatz von CAD-Systemen im Vergleich mit Einzelbefundung - liegen unter dem medianen Anteil der Sensitivitätsverbesserung für die Doppelbefundung (7,8%). Die Rate wiedereinbestellter Frauen ist bei CAD-Systemen grundsätzlich erhöht.

In Bezug auf den möglichen Schaden kann somit festgestellt werden, dass es durch den Einsatz von CAD-Systemen nicht möglich erscheint, unnötige Wiedereinbestellungen zu vermeiden, sondern diese erhöhen sich dadurch.

5. Evidenztabellen

5.1 Evidenz zu Doppelbefundung von Screening-Mammographien

5.1.1 Aggregierte Evidenz

Tabelle 1
Aggregierte Evidenz

Autor, Titel, Jahr	Dinnes J et al., 'Effectiveness and Cost-Effectiveness of Double Reading of Mammograms In Breast Cancer Screening', 2001
Schlüsselfrage	Die Screening Accuracy, das Outcome der Patientinnen und die Kosten von Doppelbefundung versus Einfachbefundung zu vergleichen. Der Fokus bei der Ergebnisdarstellung liegt allerdings auf der Rate der entdeckten Karzinome und der Recall ³ -Rate.
Einschlusskriterien	Definiert; entsprechen den hier verwendeten Einschlusskriterien
Extrahierte Studien, auf denen die Ergebnisse v.a. basieren	9 ‚vollständige‘ Kohortenstudien, davon 4 mit Follow-up (7 prospektive, 2 retrospektive) (die eingeschlossenen Studien sind in den jeweiligen Evidenztabellen mit * gekennzeichnet)
Zusammenfassendes Evidenzstatement	<ol style="list-style-type: none"> 1. Doppelbefundung erhöht die Sensitivität der Karzinomentdeckung bei Screening-Mammographien um 2,9 bis 13,7% (median 7,8%). 2. Die Spezifität kann – abhängig vom Entscheidungsverfahren nach Doppelbefundung – erniedrigt (bis zu -2,1%) oder erhöht (bis zu 2,8%) sein. Unilaterales Verfahren senkt die Spezifität , während Consensus-Verfahren oder oder Arbitrationverfahren die Spezifität eher erhöhen. 3. Kosteneffektivität im britischen National Health System (Ergebnisse aus 2 Studien) betragen die Mehrkosten der Doppelbefundung 0,12-0,69 brit. Pfund. Das Assessment für wiedereinbestellte Frauen kostet zwischen 36 und 68,67 brit. Pfund. Die Kosten pro zusätzlich entdecktem Karzinom belaufen sich auf 1,16-2,21 brit. Pfund. Am kostengünstigsten erscheint das Konsensus Verfahren.
Evidenzbewertung	2a

³ Recall = Rate wiedereinbestellter Frauen

5.1.2 Prospektive Kohortenstudien: Studiencharakteristika

Tabelle 2a

Prospektive Kohortenstudien: Studiencharakteristika

Autor/ Jahr/ Land	Studienbeschreibung		Gemessene Endpunkte
	Fokus/Intervention	Doppelbefundung im Vergleich zu Einzelbefundung	
Renaud et al., 1991, Frankreich*	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV
	Population:	n= 17, 202; 50-65 Jahre	
	Erfassungszeitraum:	1989-1990	
	Screening-Runde:	prävalent	
	Anzahl der Aufnahmen:	2, eine Ebene	
	Gesamtzahl der Befunder:	nicht bekannt	
	Erfahrung der Befunder:	nicht bekannt	
	Verblindete Befundung:	nein	
	Art der Entscheidung nach Doppelbefundung:	Arbitration-Verfahren	
Follow-up:	nein		
Antinnen et al., 1993 Finnland* ❖	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV ▪ Biopsie-Rate ▪ Sensitivität ▪ Spezifität
	Population:	n= 15, 457; 50-59 Jahre	
	Erfassungszeitraum:	1987-1988	
	Screening-Runde:	prävalent	
	Anzahl der Aufnahmen:	4, zwei Ebenen	
	Gesamtzahl der Befunder:	4	
	Erfahrung der Befunder:	nicht bekannt	
	Verblindete Befundung:	ja	
	Art der Entscheidung nach Doppelbefundung:	Konsensus-Verfahren	
Follow-up:	24 Monate		
Anderson et al., 1994, UK*	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall⁴ ▪ PPV ▪ Biopsie-Rate ▪ Tumorklassifikation ▪ Kosten
	Population:	n= 28, 170 (90% der gescreenten Frauen) 50-64 Jahre	
	Erfassungszeitraum:	7/1990-10/1991	
	Screening-Runde:	prävalent	
	Anzahl der Aufnahmen:	2, eine Ebene	
	Gesamtzahl der Befunder:	3	

⁴ Recall = Rate wiedereinbestellter Frauen

Autor/ Jahr/ Land	Studienbeschreibung		Gemessene Endpunkte
	Fokus/Intervention	Doppelbefundung im Vergleich zu Einzelbefundung	
	Erfahrung der Befunder:	3, 12, 14 Jahre	
	Verblindete Befundung:	nein	
	Art der Entscheidung nach Doppelbefundung:	Unilaterales-Verfahren	
	Follow-up:	nein	
Warren et al. 1995, UK* ❖	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV ▪ Sensitivität ▪ Spezifität ▪ Tumorklassifikation ▪ Kosten
	Population:	n= 33, 734; 50-64 Jahre	
	Erfassungszeitraum:	1987-1991	
	Screening-Runde:	prävalent	
	Anzahl der Aufnahmen:	1/5= 2, eine Ebene, sonst 4, zwei Ebenen	
	Gesamtzahl der Befunder:	10	
	Erfahrung der Befunder:	-	
	Verblindete Befundung:	ja	
	Art der Entscheidung nach Doppelbefundung:	Konsensus-Verfahren	
	Follow-up:	36 Monate	
Ciatto et al., 1995, Italien* ❖	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV ▪ Kosten
	Population:	n= 18, 817; 50-70 Jahre	
	Erfassungszeitraum:	1991-1994	
	Screening-Runde:	prävalent und inzident	
	Anzahl der Aufnahmen:	beides: 2, eine Ebene und 4, zwei Ebenen	
	Gesamtzahl der Befunder:	4	
	Erfahrung der Befunder:	-	
	Verblindete Befundung:	ja	
	Art der Entscheidung nach Doppelbefundung:	Unilaterales-Verfahren	
	Follow-up:	nein	
Pauli et al., 1996, UK* ❖	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV ▪ Sensitivität ▪ Spezifität
	Population:	n= 17, 228, 202; 50-64 Jahre	
	Erfassungszeitraum:	1995	
	Screening-Runde:	prävalent und inzident	
	Anzahl der Aufnahmen:	2, eine Ebene	
	Gesamtzahl der Befunder:	35	
	Erfahrung der Befunder:	-	
	Verblindete Befundung:	nein	
	Art der Entscheidung nach Doppelbefundung:	Unilaterales-Verfahren für Radiologen, Arbitration-Verfahren durch Zweitbefunder bei RTA als Erstbefunder	
	Follow-up:	18 Monate	

Autor/ Jahr/ Land	Studienbeschreibung		Gemessene Endpunkte
	Fokus/Intervention	Doppelbefundung im Vergleich zu Einzelbefundung	
Seradour et al., 1997, Frankreich*	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV ▪ Biopsie-Rate ▪ Sensitivität ▪ Spezifität ▪ Tumorklassifikation
	Population:	n= 258, 003; 50-64 Jahre	
	Erfassungszeitraum:	1990-1994	
	Screening-Runde:	prävalent und inzident	
	Anzahl der Aufnahmen:	2, eine Ebene	
	Gesamtzahl der Befunder:	120	
	Erfahrung der Befunder:	-	
	Verblindete Befundung:	ja	
	Art der Entscheidung nach Doppelbefundung:	Unilaterales-Verfahren	
	Follow-up:	11 Monate	
Liston, Dall, 2003, UK	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall
	Population:	n= 177, 167; 50-64 Jahre aus einer Screening-Einheit	
	Erfassungszeitraum:	4/1995-3/2002	
	Screening-Runde:	ca. 23% prävalent und 77% inzident	
	Anzahl der Aufnahmen:	4, zwei Ebenen initial	
	Gesamtzahl der Befunder:	5	
	Erfahrung der Befunder:	Zeit nicht angegeben; 5.000 Aufnahmen pro Jahr	
	Verblindete Befundung:	nein	
	Art der Entscheidung nach Doppelbefundung:	Arbitration-Verfahren	
	Follow-up:	zum Teil; für letzte Screening-Runde (2002) kein Follow-up, keine ausgewiesene Zahlen	
Fokus/Intervention		Stellenwert von Arbitration-Verfahren bei diskrepanten Befunden im Konsensus-Verfahren	
Dujim, Groene- woud, et al., 2004, Nieder- lande	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV ▪ Sensitivität ▪ Spezifität ▪ Biopsie-Rate ▪ Tumorklassifikation
	Population:	n= 65, 779; 50-75 Jahre (mittl. Alter 59 J.)	
	Erfassungszeitraum:	7/1998-1/2001	
	Screening-Runde:	17,2% prävalentes Screening, sonst inzidenten Screening	
	Anzahl der Aufnahmen:	4, zwei Ebenen initial und später nur bei Besonderheiten, ca. 30%.	
	Gesamtzahl der Befunder:	8	
	Erfahrung der Befunder:	15-36 Monate; 3.000-3.500 Aufnahmen pro Jahr	
	Verblindete Befundung:	ja	
	Art der Entscheidung nach Doppelbefundung:	Konsensus, falls Dissens: Arbitration-Verfahren durch Panel von 3 Experten	
	Follow-up:	mind. 24 Monate	

5.1.3 Retrospektive Kohortenstudien: Studiencharakteristika

Tabelle 2b
Retrospektive Kohortenstudien: Studiencharakteristika

Autor/ Jahr/ Land	Studienbeschreibung		Gemessene Endpunkte
	Fokus/Intervention	Doppelbefundung im Vergleich zu Einzelbefundung	
Thurfjell et al., 1994,# Schweden	Erhebung:	retrospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV
	Population:	n= 11, 343; 40-74 Jahre	
	Erfassungszeitraum:	1989-1990	
	Screening-Runde:	prävalent und inzident	
	Anzahl der Aufnahmen:	2, eine Ebene (ab 55Jahre, 36%) und 4, zwei Ebenen	
	Gesamtzahl der Befunder:	2	
	Erfahrung der Befunder:	-	
	Verblindete Befundung:	ja	
	Art der Entscheidung nach Doppelbefundung:	Unilaterales-Verfahren	
	Follow-up:	nein	
Williams et al., 1995*, Neusee- land	Erhebung:	retrospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV
	Population:	n= 5, 659; 50-64 Jahre	
	Erfassungszeitraum:	1991-1992	
	Screening-Runde:	prävalent	
	Anzahl der Aufnahmen:	4, zwei Ebenen	
	Gesamtzahl der Befunder:	10	
	Erfahrung der Befunder:	-	
	Verblindete Befundung:	ja	
	Art der Entscheidung nach Doppelbefundung:	Konsensus-Verfahren	
	Follow-up:	nein	
Deans et al., 1998, UK*	Erhebung:	retrospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV ▪ Kosten
	Population:	n= 258, 003; 40-64 Jahre	
	Erfassungszeitraum:	1992-1996	
	Screening-Runde:	prävalent und inzident	
	Anzahl der Aufnahmen:	2, eine Ebene (bis 7/95) und 4, zwei Ebenen	
	Gesamtzahl der Befunder:	ca. 35	
	Erfahrung der Befunder:	-	
	Verblindete Befundung:	nein	

Autor/ Jahr/ Land	Studienbeschreibung		Gemessene Endpunkte
	Fokus/Intervention	Doppelbefundung im Vergleich zu Einzelbefundung	
	Art der Entscheidung nach Doppelbefundung:	Unilaterales-Verfahren	
	Follow-up:	nein	
Kwek et al., 2003 , Singapore	Erhebung:	retrospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV ▪ Biopsie-Rate ▪ PPV Biopsie
	Population:	n= 28, 231; 50-64 Jahre	
	Erfassungszeitraum:	30 Monate (1994-1996)	
	Screening-Runde:	prävalent und geringer Anteil inzident	
	Anzahl der Aufnahmen:	4, zwei Ebenen	
	Gesamtzahl der Befunder:	6	
	Erfahrung der Befunder:	alles erfahrene Befunder	
	Verblindete Befundung:	ja	
	Art der Entscheidung nach Doppelbefundung:	Unilaterales-Verfahren	
	Follow-up:	nein	
Ciatto et al., 2005 , Italien	Erhebung:	retrospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall⁵ ▪ PPV
	Population:	n= 177, 631; 50-69 Jahre	
	Erfassungszeitraum:	1/1998-6/2003	
	Screening-Runde:	vorwiegend inzident	
	Anzahl der Aufnahmen:	2, eine Ebene bei inzidentem Screening	
	Gesamtzahl der Befunder:	11	
	Erfahrung der Befunder:	alle seit Jahren spezialisiert	
	Verblindete Befundung:	nein	
	Art der Entscheidung nach Doppelbefundung:	Unilaterales-Verfahren	
	Follow-up:	nein	

5.1.4 Prospektive Kohortenstudien - Ergebnisse

Tabelle 3a

Prospektive Kohortenstudien - Ergebnisse

⁵ Recall = Rate wiedereinbestellter Frauen

Autor, Jahr, Land, Gesamtzahl, Entscheidung bei diskrepanten Befunden	Ergebnisse: Einzelbefundung Doppelbefundung		Bemerkungen	Schlussfolgerungen/ Empfehlungen	Evidenz-grad
Renaud et al., 1991 , Frankreich* n= 17,228 Arbitration	<i>Rate entdeckter Karzinome:</i> Bef. 1 0,35% Bef. 2 0,43% <i>Recall</i> Bef. 1 9,5% Bef. 2 7,7% <i>PPV:</i> Bef. 1 0,036% Bef. 2 0,055%	0,44% 7,7% 0,057 %	Pilotprojekt mit niedergelassenen Radiologen, keine Spezialisten (Recall hoch) 16 zusätzliche Cas bei Bef. 1, 2 zusätzliche Cas bei Bef. 2	Doppelbefundung wird empfohlen	3b
Antinnen et al., 1993 , Finnland* n= 15,457 Konsensus ❖	<i>Rate entdeckter Karzinome:</i> Bef. 1 0,41% Bef. 2 0,39% <i>Recall:</i> Bef. 1 3,8% Bef. 2 2,8% <i>PPV:</i> Bef. 1 0,11 Bef. 2 0,14 NPV: Bef. 1 0,97 Bef. 2 0,94 Sensitivität: Bef. 1 94,3% Bef. 2 87,1% Spezifität: Bef. 1 96,7% Bef. 2 97,6% Biopsie-Rate: - PPV Biopsie: -	0,44% 2,5% 0,17 0,99 97,1% 97,9% 0,7% 0,66	Sensitivität für Bef. 1 um 2,9% erhöht, für Reader 2 um 10%! Spezifität bei beiden geringgradig erhöht	Unabhängige Doppelbefundung mit Konsensus-Verfahren wird empfohlen	2b

Autor, Jahr, Land, Gesamtzahl, Entscheidung bei diskrepanten Befunden	Ergebnisse:		Bemerkungen	Schlussfolgerungen/ Empfehlungen	Evidenz- grad	
	Einzelbefundung	Doppelbefundung				
Anderson et al., 1994 , UK* n= 28,170 Unilateral	<i>Rate entdeckter Karzinome:</i>	0,61%	Doppelbefundung nicht verblindet 21 (10,9%) Cas wurden von einem der beiden Befunder nicht erkannt, also zusätzlich durch unilateralen Recall entdeckt	Doppelbefundung wird unter der Annahme einer erhöhten Sensitivität empfohlen	3b	
	Bef. 1					0,61%
	Bef. 2					0,59%
	Bef. 3	0,49%				
<i>Recall⁶:</i>	6,0%					
Bef. 1					5,1%	
Bef. 2					7,8%	
Bef. 3	4,0%					
<i>PPV:</i>	0,10					
Bef. 1					0,12%	
Bef. 2					0,07%	
Bef. 3	0,12%					
<i>Biopsie-Rate:</i>	0,8%					
Bef. 1					0,9%	
Bef. 2					0,8%	
Bef. 3	0,7%					
Ciatto et al., 1995 , Italien* n= 18,117 Unilateral	<i>Rate entdeckter Karzinome:</i>	0,66%		Bei nur gering erhöhter Anzahl an mehr erkannten Karzinomen und einer höheren Recall-Rate wird die Efficiency der Doppelbefundung in Frage gestellt	3b	
	Bef. 1					0,62%
	Bef. 2	0,65%				
	<i>Recall:</i>	4,0%				
Bef. 1	3,3%					
Bef. 2	3,6%					
<i>PPV:</i>	0,16					
Bef. 1					0,18	
Bef. 2	0,18					

⁶ Recall = Rate wiedereinbestellter Frauen

Autor, Jahr, Land, Gesamtzahl, Entscheidung bei diskrepanten Befunden	Ergebnisse: Einzelbefundung Doppelbefundung		Bemerkungen	Schlussfolgerungen/ Empfehlungen	Evidenz-grad
Pauli et al., 1996, UK* n=17,202 Konsensus +Arbitration ❖	<i>Rate entdeckter Karzinome:</i> Bef. 1 0,61% Bef. 2 0,47% Bef. 3 0,60%	0,67%	Bef. 1 und Bef. 2 sind RTAs	Doppelbefundung – auch mit RTAs – wird empfohlen	3b
<i>Recall:</i> Bef. 1 7,7% Bef. 2 6,6% Bef. 3 5,4%	5,6%				
<i>PPV:</i> n. ermittelbar (absolute Zahlen fehlen)	n. ermittelbar				
<i>Sensitivität:</i> Bef. 1+Bef. 2 73% Bef. 3 73%	79,3%				
<i>Spezifität:</i> Bef. 1 und Bef. 2 277% Bef. 3 84%	94%				
Warren et al., 1995,UK* n= 33.734 Konsensus und Arbitration ❖	<i>Rate entdeckter Karzinome:</i> Bef. 1 0,69%	0,79%%	Rate entdeckter Karzinome von Reader 2 fehlt Unilateraler Recall wäre 10% gewesen, bei +3 entdeckten Cas, die aufgrund des Konsensus nicht entdeckt wurden Sensitivität und Spezifität signifikant höher (Mc Nemar) Sensitivität der Befundung steigerte sich über den Zeitraum	Doppelbefundung mit Konsensus-Verfahren bzw. Arbitration bei weiterbestehender Uneinigkeit wird vor allem in den ersten Jahren eines Screening-Programms empfohlen.	2b
<i>Recall:</i> Bef. 1 6,9% Bef. 2 6,5%	4,2%%				
<i>PPV:</i> Bef. 1 0,10%	0,19%				
<i>NPV</i> 0,67%	0,76%				
<i>Sensitivität:</i> Bef. 1 82,8%	93,1%				
<i>Spezifität:</i> Bef. 1 93,8%	96,6%				

Autor, Jahr, Land, Gesamtzahl, Entscheidung bei diskrepanten Befunden	Ergebnisse:		Bemerkungen	Schlussfolgerungen/ Empfehlungen	Evidenz-grad	
	Einzelbefundung	Doppelbefundung				
Seradour et al., 1997 , Frankreich* n= 258,003 Unilateral ❖	<i>Rate entdeckter Karzinome:</i> Bef. 1	0,53%	Dezentrales Programm, Bef. 2 ist sehr viel erfahrener	Doppelbefundung wird aufgrund der höheren Sensitivität empfohlen	2b	
	Bef. 2					0,45%
	<i>Recall:</i>	4,5%				
	Bef. 1					3,6%
	Bef. 2	2,4%				
	<i>PPV:</i>	0,11%				
	Bef. 1					0,12%
	Bef. 2	0,19%				
<i>NPV:</i>						
Bef. 1		0,85%				
Bef. 2	0,83%					
Sensitivität:	88,9%					
Bef. 1		75,2%				
Bef. 2	79,4%					
Spezifität:	96,0%					
Bef. 1		96,8%				
Bef. 2	98,0%					
Biopsie-Rate:	0,92%					
Liston, Dall, 2003 , UK n= 177,167 Arbitration	<i>Rate entdeckter Karzinome:</i>	0,60%	Durch Doppelbefundung wurden 87 Karzinome mehr erkannt (+8,1%)	Doppelbefundung wird empfohlen	3b	
	Bef. 1					0,55%
	Bef. 2	0,60%				
<i>Recall:</i>	5,0%	-				
<i>PPV:</i> kann nicht angegeben werden (absolute Zahlen fehlen)	kann nicht angegeben werden (absolute Zahlen fehlen)					
Ergebnisse: Unilateral Konsensus						
Dujim, Groenewoud, et al.,	<i>Rate entdeckter Karzinome</i>	0,49%	0,48%	Usus in den Niederlanden ist Doppelbefundung mit	Arbitration bei diskrepantem Befund erscheint nicht	2b

Autor, Jahr, Land, Gesamtzahl, Entscheidung bei diskrepanten Befunden	Ergebnisse:				Bemerkungen	Schlussfolgerungen/ Empfehlungen	Evidenz-grad
	Einzelbefundung	Doppelbefundung					
2004, Niederlande n= 65,779 Konsensus und Arbitration, falls Konsensus nicht erfolgreich ❖	Recall:	1,1%		0,09%	Konsensus-Verfahren. Jetzt: Arbitration mit speziellem Expertenpanel mit 3 Experten, jede suspekta Befundung führt zum Recall. 3 Karzinome aufgrund von Arbitration nicht erkannt	sinnvoll, da Karzinome übersehen werden können. Bei diskrepanten Befunden, über die sich Erst- und Zweitbefunder nicht einigen können, sollte ein Recall erfolgen, grundsätzlich sollte das Konsensus-Verfahren beibehalten werden.	
	PPV:	0,57		0,5			
	Sensitivität	76,6%	76,3%	75,6%			
	Spezifität	-	-	99,5%			
	NPV	0,99		0,99			

5.1.5 Retrospektive Kohortenstudien- Ergebnisse

Tabelle 3b
Retrospektive Kohortenstudien – Ergebnisse

Autor, Jahr, Land, Gesamtzahl, Entscheidung bei diskrepanten Befunden	Ergebnisse:				Bemerkungen	Schlussfolgerungen/ Empfehlungen	Evidenz-grad
	Einzelbefundung	Doppelbefundung					
Thurfjell et al., # 1994, Schweden, n= 11,343	<i>Rate entdeckter Karzinome:</i> Bef. 1 0,61 % Bef. 2 0,54% Recall: - PPV: 0.6		0,68%	-	Vergleichende Daten zu Recall liegen nicht vor	Wird empfohlen	3b
			0.6				
Williams et al., 1995* , Neuseeland	<i>Rate entdeckter Karzinome:</i> Bef. 1 0,99% Bef. 2 0,97%		1,02%		Doppelbefundung mit Konsensus wird empfohlen	3b	

Autor, Jahr, Land, Gesamtzahl, Entscheidung bei diskrepanten Befunden	Ergebnisse: Einzelbefundung Doppelbefundung	Bemerkungen	Schlussfolgerungen/ Empfehlungen	Evidenz-grad	
n=5,659 Konsensus	<i>Recall:</i> Bef. 1 9,4% Bef. 2 12,4%	10,3%			
Deans et al., 1998 , UK* n= 258,003	<i>Rate entdeckter Karzinome:</i> Bef. 1 0,47% Bef. 2 0,52% <i>Recall:</i> Bef. 1 4,8% Bef. 2 5,2% <i>PPV:</i> Bef. 1 0,097 Bef. 2 0,089	0,53%		Doppelbefundung wird empfohlen	3b
Kwek et al., 2003 , Singapore, n= 28,231 Unilateral	<i>Rate entdeckter Karzinome:</i> Bef. 1 (130) 0,46% Bef. 2 (126) 0,44% <i>Recall:</i> Bef. 1 (1669) 5,91% Bef. 2 (1437) 5,09% <i>PPV:</i> 0,0832 <i>Biopsie-Rate:</i> Bef. 1 (300) 1,06% Bef. 2 (297) 1,05% <i>PPV Biopsie:</i> Bef. 1 43,2% Bef. 2 42,4%	(135) 0,47% (2185) 7,7% 0,061 (329) 1,17% 41,0%		Doppelbefundung mit unilateralem Recall wird empfohlen	3b
Ciatto et al., 2005 , Italien n= 177,631 Unilateral	Rate entdeckter Karzinome: Bef. 1 0,37% Bef. 2 0,39% Recall ⁷ : Bef. 1 2,98% Bef. 2 3,15%	0,40% 3,59	43 Cas zusätzlich durch second reading erkannt	Der Nutzen von unilateralem Doppelbefundung als gering aber vertretbar bei gering erhöhter Recall-Rate gewertet	3b

⁷ Recall = Rate wiedereinbestellter Frauen

Autor, Jahr, Land, Gesamtzahl, Entscheidung bei diskrepanzen Befunden	Ergebnisse: Einzelbefundung Doppelbefundung	Bemerkungen	Schlussfolgerungen/ Empfehlungen	Evidenz -grad
	<i>PPV:</i> Bef. 1 0,13 Bef. 2 0,12	0,11		

5.2 Evidenz zum Einsatz von CAD-Systemen beim Mammographie-Screening

5.2.1 Kohortenstudien: Studiencharakteristika

Tabelle 4

Kohortenstudien: Studiencharakteristika

Autor/ Jahr/ Land	Studienbeschreibung		Gemessene Endpunkte
	Fokus/Intervention	Einzelbefundung mit CAD im Vergleich zu Doppelbefundung	
Khoo et al., 2005, UK	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ PPV ▪ Größe und Invasivität der entdeckten Tumore ▪ Zeit für Befundung
	Population:	n= 6.111; 45-94 (mittl. Alter 58,4)	
	Erfassungszeitraum:	3/2003-9/2004	
	Screening-Runde:	nicht bekannt	
	Anzahl der Aufnahmen:	4, zwei Ebenen initial	
	Gesamtzahl der Befunder:	12	
	Erfahrung der Befunder:	erfahrene Befunder, 5 RTAs, alle mind. 5.000 Aufnahmen pro Jahr	
	Verblindete Befundung:	ja	
	Art der Entscheidung nach Doppelbefundung:	Arbitration-Verfahren	
	Angabe der Einführung in CAD	nein	
Gilbert et al., 2006, UK	Erhebung:	retrospektiv (Doppelbefundung) und prospektiv (Einzelbefundung mit CAD)	<ul style="list-style-type: none"> ▪ PPV ▪ Sensitivität ▪ Spezifität
	Population:	n= 10.267; > 50 Jahre	
	Erfassungszeitraum:	1996	
	Screening-Runde:	nicht bekannt	
	Anzahl der Aufnahmen:	4, zwei Ebenen initial	
	Gesamtzahl der Befunder:	8	
	Erfahrung der Befunder:	alle mind, 5.000 Aufnahmen pro Jahr	
	Verblindete Befundung:	ja	
	Art der Entscheidung nach Doppelbefundung:	Unilateral und Arbitration-Verfahren (2 Zentren)	
	Angabe der Einführung in CAD:	2 Monate Training, zunächst mit 25% Cas, dann mit 5%	
Follow-up:	ja, nicht vollständig		
Follow-up:	nicht vollständig, nicht auswertbar		
Fokus/Intervention		Einzelbefundung im Vergleich zu Einzelbefundung mit CAD	

Autor/ Jahr/ Land	Studienbeschreibung		Gemessene Endpunkte
	Fokus/Intervention	Einzelbefundung mit CAD im Vergleich zu Doppelbefundung	
Freer et al. 2001, USA	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV ▪ Biopsierate ▪ PPV für Biopsie
	Fokus:	Einzelbefundung mit CAD im Vergleich zu Einzelbefundung	
	Population:	n= 12.860, 22-88Jahre, median 49 J.	
	Erfassungszeitraum:	2/1999-2/2000	
	Screening-Runde:	27% inzident	
	Anzahl der Aufnahmen:	4, zwei Ebenen	
	Gesamtzahl der Befunder:	2	
	Erfahrung der Befunder:	erfahrene Befunder mit spezieller Ausbildung	
	Verblindete Befundung:	ja	
	Angabe der Einführung in CAD:	ja (Befundung von 2.500 Aufnahmen vor der Studie)	
Follow-up:	nein		
Helvie et al., 2004, USA	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV ▪ Biopsierate ▪ PPV für Biopsie
	Fokus:	Einzelbefundung mit CAD im Vergleich zu Einzelbefundung	
	Population:	n= 2.389 (Studiengruppe aus Gesamtpopulation) 26-85Jahre, median 52 J.	
	Erfassungszeitraum:	-	
	Screening-Runde:	inzident	
	Anzahl der Aufnahmen:	4, zwei Ebenen	
	Gesamtzahl der Befunder:	13	
	Erfahrung der Befunder:	alle spezialisiert zwischen 3 und 24 J.	
	Verblindete Befundung:	ja	
	Angabe der Einführung in CAD:	nein	
Follow-up:	nur für 1.000 Fälle		
Cupples et al., 2005, USA	Erhebung:	retrospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall⁸ ▪ PPV ▪ Biopsierate ▪ PPV für Biopsie
	Fokus:	Einzelbefundung mit CAD im Vergleich zu Einzelbefundung	
	Population:	n= 27.274, 50-74 Jahre	
	Erfassungszeitraum:	1/1998-12/2000	
	Screening-Runde:	prävalent und inzident	
	Anzahl der Aufnahmen:	4, zwei Ebenen	
	Gesamtzahl der Befunder:	4	
	Erfahrung der Befunder:	alle spezialisiert, 1 Befunder befundete nur mit CAD	
	Verblindete Befundung:	ja	

⁸ Recall = Rate wiedereinbestellter Frauen

Autor/ Jahr/ Land	Studienbeschreibung		Gemessene Endpunkte
	Fokus/Intervention	Einzelbefundung mit CAD im Vergleich zu Doppelbefundung	
	Angabe der Einführung in CAD:	nein	
	Follow-up:	nein	
Dean et al., 2006, USA	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV
	Fokus:	Einzelbefundung mit CAD im Vergleich zu Einzelbefundung	
	Population:	n= 9.520, 50-74 Jahre	
	Erfassungszeitraum:	28 Monate	
	Screening-Runde:	inzident	
	Anzahl der Aufnahmen:	nicht bekannt	
	Gesamtzahl der Befunder:	1	
	Erfahrung der Befunder:	sehr erfahren	
	Verblindete Befundung:	ja	
	Angabe der Einführung in CAD:	ja	
	Follow-up:	nein	
Ko et al., 2006, USA	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ ▪ ▪ Recall ▪ PPV ▪ Biopsie-Rate ▪ PPV Biopsie ▪ Sensitivität ▪ Spezifität
	Fokus:	Einzelbefundung mit CAD im Vergleich zu Einzelbefundung	
	Population:	n= 5.016, 50-74 Jahre	
	Erfassungszeitraum:	26 Monate	
	Screening-Runde:	inzident	
	Anzahl der Aufnahmen:	nicht bekannt	
	Gesamtzahl der Befunder:	2	
	Erfahrung der Befunder:	erfahrene Radiologen	
	Verblindete Befundung:	ja	
	Angabe der Einführung in CAD:	ja	
	Follow-up:	ja	
Mortonet al., 2006, USA	Erhebung:	prospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall ▪ PPV ▪ Biopsie-Rate ▪ PPV für Biopsie
	Fokus:	Einzelbefundung mit CAD im Vergleich zu Einzelbefundung	
	Population:	n= 18.096, 23-98 Jahre, median 60 Jahre	
	Erfassungszeitraum:	4/2001-10/2002	
	Screening-Runde:	prävalent und inzident	
	Anzahl der Aufnahmen:	nicht bekannt	
	Gesamtzahl der Befunder:	12	
	Erfahrung der Befunder:	Spezialisten, 3-30 Jahre Erfahrung	
	Verblindete Befundung:	nein	
Angabe der Einführung in CAD:	nein		

Autor/ Jahr/ Land	Studienbeschreibung		Gemessene Endpunkte
	Fokus/Intervention	Einzelbefundung mit CAD im Vergleich zu Doppelbefundung	
	Follow-up:	nein	
Fenton et al., 2007, USA	Erhebung:	retrospektiv	<ul style="list-style-type: none"> ▪ Rate entdeckter Karzinome ▪ Recall⁹ ▪ PPV ▪ Sensitivität ▪ Spezifität ▪ Biopsie-Rate ▪ PPV-Biopsie
	Fokus:	Einzelbefundung mit CAD im Vergleich zu Einzelbefundung in 7 Einrichtungen	
	Population:	n= 2.22,1.355 > 40 J.	
	Erfassungszeitraum:	1998-2002	
	Screening-Runde:	prävalent und inzident	
	Anzahl der Aufnahmen:	nicht bekannt	
	Gesamtzahl der Befunder:	n= 122	
	Erfahrung der Befunder:	erfahrene Radiologen, weniger Erfahrung bei Einrichtungen mit CAD	
	Verblindete Befundung:	-	
	Angabe der Einführung in CAD:	-	
	Follow-up:	mind. 12 Monate (1 Screening-Intervall)	

⁹ Recall = Rate wiedereinbestellter Frauen

5.2.2 Kohortenstudien: Ergebnisse

Tabelle 4
Kohortenstudien: Ergebnisse zu CAD

Autor, Jahr, Land,	Ergebnisse		Bemerkungen	Schlussfolgerungen/ Empfehlungen	Evidenz- grad
	Doppelbefundung	Einzelbefundung mit CAD			
Khoo, Taylor, Given-Wilson 2005 , UK n= 6.111	<i>Rate entdeckter Karzinome:</i> (61) 0,99%	(54) 0,89%	2 Karzinome nur durch CAD, von 9 angezeigten richtigen Prompts werden 7 durch die Befunder nicht ausgewählt. CAD zeigte 84% der Ca insgesamt richtig an. ,Falsche Prompts': 1,59 pro Fall. Zeit für Einzelbefundung: 25s Zeit für Einzelbefundung +CAD: 45 s	Durch Doppelbefundung werden mehr Karzinome erkannt als durch Einzelbefundung mit CAD. 78% der richtig markierten CAD-Prompts wurden von den Befundern nicht ausgewählt. CAD nimmt fast 50% mehr Zeit in Anspruch als Einzelbefundung	3b
	<i>Recall:</i> (599) 9,8%	(635) 10,5%			
	<i>PPV:</i> 0,10	0,085			
Gilbert, Astley, McGee, et al., 2006 , UK n= 10.096	<i>Rate entdeckter Karzinome:</i> 0,97%	1,11%		Einzelbefundung mit CAD wird als mindestens gleichgut wie Doppelbefundung gewertet. Eine prospektive randomisierte Studie dazu wird empfohlen	3b
	<i>Recall¹⁰:</i> (661) 6,5%	(869) 8,6%			
	<i>PPV:</i> 0,14	0,13			
	Ergebnisse	Einzelbefundung			
	Einzelbefundung	mit CAD			
Freer, Ulissey, 2001 , USA	<i>Rate entdeckter Karzinome:</i> 41 0,31%	49 0,38%	Pro 4 Aufnahmen, zwei Ebenen Mammographie wurden im Mittel 2,8 CAD- Prompts gesetzt, davon	CAD erscheint hilfreich für die verbesserte Ca- Entdeckung. Weitere Studien erforderlich	3b

¹⁰ Recall = Rate wiedereinbestellter Frauen

		Ergebnisse			
		Einzelbefundung	Einzelbefundung mit CAD		
n=12.860		<i>Recall:</i> 830 6,45%	986 7,66%	wurden 2,6% als suspekt bewertet, 97,4% wurden verworfen!	
		<i>PPV:</i> 0,049	0,049		
Helvie, Hadjiiski, et al., 2004 , USA		<i>Rate entdeckter Karzinome:</i> 10 0,41%	11 0,46%	Die Recall-Rate in der Studiengruppe war um 30% erhöht: Nicht-Studiengruppe hatte 10,3% Recall CAD markierte 10 von 11 CAs	3b
n= 2.389		<i>Recall:</i> 344 14,3%	378 15,8%		
		<i>PPV:</i> 0,029	0,029		
Cupples, Cunnigham, et al., 2005 , USA		<i>Rate entdeckter Karzinome:</i> 0,37%	0,43%	Entdeckung kleiner invasiver CA (< 1 cm) erhöhte sich um 164% P=0.069	3b
n= 27.274		<i>Recall:</i> 7,7%	8,1%		
		<i>PPV:</i> 0,049	0,05	Ergebnisse sind alle nicht statistisch signifikant	
		Biopsie-Rate: 1,4%	1,5%		
		<i>PPV Biopsie:</i> 26,9%	28,6%		
Dean, Ilvento, 2006 , USA		<i>Rate entdeckter Karzinome:</i> 0,53%	0,6% (+4Ca)	Biopsie-Rate nicht für Screening-Patientinnen ausgewiesen	3b
n=5.631 (Screening-Gruppe von ges. 9.520)		<i>Recall:</i> 6,2%	7,8% (p=0,04)		
		<i>PPV:</i> 0,08	0,077	Durch CAD entdeckte Karzinome kleiner als durch Radiologen entdeckte, Unterschied nicht statistisch signifikant	
Ko, Nicholas, et al., 2006 , USA		<i>Rate entdeckter Karzinome:</i> 0,79%	0,89%	1 durch CAD markiertes Karzinom wurde vom Radiologen verworfen	2b
n=5.016		<i>Recall:</i> 12%	14%		
		<i>PPV:</i> 0,07	0,064	CAD erhöht signifikant die Recall-Rate und kann die Sensitivität um mind. 4% erhöhen	
		<i>PPV-Biopsie:</i> 36,8%	36,3%		
		Sensitivität: (40/48) 90%	(45/48) 94%		
		Spezifität: 99%	99%		

		Ergebnisse				
		Einzelbefundung	Einzelbefundung mit CAD			
Morton, Whaley, 2006 , USA n= 18.096	<i>Rate entdeckter Karzinome:</i>	105 0,58%	113 0,62%	Studie bezieht Interventionen auf Anzahl der Mammographien, hier dargestellt sind Interventionen bezogen auf die Anzahl der gescreenten Frauen	CAD wird empfohlen	3b
	<i>Recall:</i>	2101 11,6%	2300 12,7%			
	<i>PPV:</i>	0,049	0,049			
	<i>Biopsie-Rate:</i>	256 1,41%	277 1,53%			
	<i>PPV –Biopsie:</i>	41,0%	40,7%			
Fenton, et al., 2007 , USA	<i>Rate entdeckter Karzinome:</i>	0,415%	0,420%	Vergleich mit weiteren Einrichtungen, die CAD nie eingeführt haben nicht dargestellt	CAD erhöht die Sensitivität nicht signifikant, aber die Spezifität und der PPV verschlechtern sich signifikant, ebenso erhöht sich die Biopsie-Rate signifikant	2b
	<i>Recall:</i>	10,1%	13,2% (<i>p</i> <0,001)			
	<i>PPV:</i>	0,41	0,32 <i>p</i> =0,01			
	<i>Biopsie-Rate:</i>	14,7%	17,6% (<i>p</i> <0,001)			
	<i>Sensitivität:</i>	80,4%	84,0% <i>n.s.</i>			
	<i>Spezifität:</i>	90,2%	87,2% <i>p</i> <0,001	Weitere Studien mit sehr großer Fallzahl gefordert		

5.3 Ausgeschlossene Studien

5.3.1 Ausgeschlossene Studien zur Doppelbefundung

Tabelle 5
Ausgeschlossene Studien zur Doppelbefundung

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Thurfjell et al., # 1994, Schweden	Fall-Kontroll-Studie n= 120 Kein Follow-up	Kein Mammographie-Setting
Beam et al., # 1996, USA	Fall-Kontroll-Studie n=79 Kein Follow-up	Kein Mammographie-Setting
Denton und Field,# 1997, UK	Retrospektive Kohortenstudie, 62,5% mit Doppelbefundung	Keine durchgängiges Doppelbefundung
Blanks et al., # 1998, UK	Retrospektive Kohortenstudie Recall unilateral, consensus und arbitration und Mix, kein Follow-up	Retrospektiv und Angaben zur Gesamtpopulation fehlen, Recall-Policy zum Teil nicht zuordenbar
Ciatto et al., 2005, Italien	Retrospektive Fallstudie, Arbitration-Verfahren, Follow-up nur zum Teil	Retrospektiv, keine vollständige Kohorte
Groenewoud et al., 2006, Niederlande	Fall-Kontroll-Studie, n= 500, fokussiert auf Kosten-Effektivität des Entscheidungsverfahrens	Kein Mammographie-Setting
Hukkinen et al., 2006, Finnland	Fall-Kontroll-Studie, n= 200, fokussiert auf Anzahl der Befunder (bis zu 6)	Kein Mammographie-Setting

5.3.2 Ausgeschlossene Studien zur Befundung mit CAD-Systemen

Tabelle 6
Ausgeschlossene Studien zur Befundung mit CAD-Systemen

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Kegelmeyer, Pruneda et al., 1994, USA	Retrospektive Fall-Kontroll-Studie n= 85, Beurteilung unregelmäßig begrenzter Herde ohne und mit CAD	Hohe Karzinomprävalenz bei kleiner Fallzahl, kein reales Screening-Setting
Thurfjell, Thurfjell et al., 1998, Finnland	Retrospektive Fall-Kontroll-Studie n= 120 zur Beurteilung ohne und mit CAD	Kein Screening-Setting
Warren Burhenne et al., 2000, USA	Retrospektive Fall-Studie, n= 427, Verblindete Beurteilung von Voraufnahmen von durch folgendes Screening entdeckten Karzinomen ohne und mit CAD	Für retrospektive Analyse rel. hohe Fallzahl, jedoch kein reales Screening-Setting
Moberg, Bjurstam, 2001, Schweden	Retrospektive Fall-Kontroll-Studien n= 59 Intervallkarzinome gemischt mit normalen Mammographien im Verhältnis 1:5, Beurteilung mit und ohne CAD, Ca-Entdeckung der Befunder 2x erhöht, 1x erniedrigt	Karzinomprävalenz 16,6% Keine reale Screening-Situation
Baker, Rosen, 2003, USA	Retrospektive Beurteilung von im Screening entdeckten Architekturstörungen durch Radiologen und durch CAD n= 45, zeigt schlechte Sensitivität von CAD für Architekturstörungen	Kleine Fallzahl, kein Screening-Setting
Brem, Baum, 2003, USA	Retrospektive Beurteilung von n= 177 ‚Missed‘ Screening-Mammographien, bei denen nach Zusatzaufnahmen eine Normalbefundung vorgenommen worden war. Evaluation durch CAD. Simuliertes Einzelbefundung mit CAD	Kein reales Screening-Setting
Ciatto, Brancato, 2003, Italien	Retrospektive Fall-Kontroll-Studie zur Beurteilung von durch Screening entdeckten Karzinomen, Simulierte Doppelbefundung und Einzelbefundung mit CAD	Kein reales Screening-Setting
Ciatto, RosselideITurCo,	Retrospektive Fall-Kontrollstudie n= 120 zur Beurteilung von Intervall-Karzinomen mit und ohne	Kein reales Screening-Setting

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
2003, Italien	CAD, simulierte Doppelbefundung durch Kombinationen von Einzelbefundung	
Karssemeijer et al., 2003, Niederlande	Retrospektive Fall-Kontroll-Studie zu Doppelbefundung und Einzelbefundung mit CAD n= 500, bezieht sich nur auf Herdbefunde und Intervallkarzinome, Doppelbefundung bringt Verbesserung um 10,5% CAD um 7% im Vergleich zu Einzelbefundung	Karzinomprävalenz >25%, keine reale Screening-Situation
Destounis et al., 2004, USA	Retrospektive Fallstudie n= 318 Voraufnahmen von durch folgendes Screening entdeckten Karzinomen ohne und mit CAD zeigt Verbesserung der FN um fast ein Drittel durch CAD	Für eine retrospektive Studie rel. hohe Fallzahl, jedoch keine reale Screening-Situation
Taylor et al., 2004, UK	HTA-Bericht, 2 konsekutive Fall-Kontroll-Studien Beurteilung ohne und mit CAD	Jeweils hohe Karzinomprävalenz bei geringer Fallzahl, kein Screening-Setting
Birdwell et al., 2005, USA	Retrospektive Fallstudie n= 286 Re-Beurteilung von Retrospektiv sichtbaren Karzinomen ohne und mit CAD	Kein Screening-Setting, ausschließlich retrospektiv identifizierte Karzinome
Hukkinen, Pamilo, 2005, Finnland	Retrospektive Beurteilung von n= 69 durch Screening entdeckte Karzinome und Voraufnahmen,	Sehr geringe Fallzahl, hohe Karzinomprävalenz, kein Screening-Setting
Menna, Di Virgilio, 2005, Italien	Retrospektive Fall-Studie, n= 98, durch unabhängige Doppelbefundung entdeckte Karzinome wurden durch ein CAD-System beurteilt, das nicht alle Karzinome identifizierte	Retrospektiv, geringe Fallzahl 100% Karzinomprävalenz, Kein Screening-Setting, kein direkter Vergleich von Einzelbefundung mit CAD und Doppelbefundung
Bornefalk, 2006, Schweden	Theoretische Berechnung der Effizienz von CAD als second reader (nach Bayes)	Modellrechnung, keine realen Daten
Ciatto et al., 2006, Italien	Retrospektive Fall-Kontroll-Studie N= 95, 75 Kontrollen Einzelbefundung und Einzelbefundung mit CAD	25% Karzinome, sehr geringe Fallzahl, kein Screening-Setting
Frigas, Kapsimalagou, 2006, Griechenland	Retrospektive Fall-Kontroll-Studie, n= 200, Beurteilung ohne und mit CAD	Als Fall-Kontroll-Studie hohe Karzinomprävalenz, deshalb kein Screening-Setting
Hukkinen, Vemas, et al., 2006, Finnland	Retrospektive Fall-Kontrollstudie n= 200, 51 Karzinome, Beurteilung durch unterschiedlich erfahrene Befunder, keine statistisch signifikanten Unterschiede ohne und mit CAD in der Karzinom-Entdeckung	25% Karzinome, Fokus auf Bewertung der Erfahrung der Befunder, kein Screening-Setting
Lindfors, McGahan et al., 2006, USA	Kosten-Effektivitätsstudie, Markov-Modell mit Annahmen verschiedener CA-Entdeckungsraten für CAD	Theoretische Berechnung, keine realen Daten
Pai, Gregory, 2006, USA	Retrospektive Fall-Studie, n=55, Beurteilung von durch Screening-Mammographie entdeckte DCIS-Fälle und die Voraufnahmen ohne und mit CAD	Sehr geringe Fallzahl, retrospektive Beurteilung, kein Screening-Setting
Taplin, Rutter, 2006, USA	Retrospektive Fall-Kontroll-Studie n= 341	Kein Screening-Setting
Timp, Karssemeijer, 2006, Niederlande	Experimentelle Fall-Kontroll Studie n= 589/938 Vergleich Rate entdeckter Karzinome zwischen a) mit Vorbefund verlinkter CAD Markierung und b) nicht verlinkter CAD Markierung, a) ergab ein besseres Ergebnis.	Kein Screening-Setting Fokus auf Unterschied Verlinkung mit Vorbefund
Skaane, Khirsagar, 2007, USA	Retrospektive Fall-Studie, n= 55, Beurteilung Screening-(Vor)Aufnahmen von Intervall-Cas und Cas der nächsten Runde ohne und mit CAD	Isolierte retrospektive Beurteilung der 55 Aufnahmen, deshalb kein Screening-Setting

6. Kosten-Nutzen-Analysen

Tabelle 7
Kosten-Nutzen-Analysen

Autor, Jahr, Land	Inhalt	Ergebnis
Anderson et al., 1994, Schottland	Zusatz-Kosten der unilateralen Doppelbefundung gesamt in einem Screening-Setting in Schottland wurden berechnet.	Direkte Kosten: 4.373 engl. Pfund Indirekte Kosten: 9.400 engl Pfund Gesamt: 13.773 engl. Pfund Pro gescreenter Frau: 0,44 engl. Pfund (ca. 0,7 €) Doppelbefundung wird als kosteneffektiv erachtet
Ciatto et al., 1995, Italien	Berechnung der zusätzl. Kosten von unabhängiger unilateraler Doppelbefundung anhand realer Ausgaben	Kosten pro Frau : 3,5 US\$ bei erster Runde (+8,5%) und 2,3 US\$ (+6,2%) bei nächster Runde Kosten pro entdecktem Karzinom: 196 US\$ (3,5%) bei erster Runde, 251 US\$ (2,7%) bei nächster Runde zusätzl. entdeckte Karzinome durch Doppelbefundung: erste Runde 4,8%, nächste Runde: 3,4% Anstieg der Wiedereinbestellung: 15,0% und 16,4% Sehen Kosten-Effektivität von Doppelbefundung kritisch
Brown et al., 1996, England	Beurteilung von drei Befundungsstrategien: Einzelbefundung, unilaterales Doppelbefundung und Doppelbefundung mit Konsensus anhand von Mammographien von 33.734 Frauen und 132 Wiedereinbestellten zur Kalkulierung der Kosten.	Durch unilaterale Doppelbefundung wurden im Vergleich zu Einzelbefundung 10 zusätzl. Karzinome pro 100.000 Frauen gefunden bei höherer Wiedereinbestellrate, durch DR mit Konsensus 9 zusätzl. Karzinome, bei signifikant niedrigerer Wiedereinbestellungsrate im Vergleich zu Einzelbefundung . Unilaterale Doppelbefundung war teurer als Einzelbefundung (+19259 engl. Pfund pro 100.000 Frauen, ca. + 0,3 € pro Frau), Doppelbefundung mit Konsensus war billiger als Einzelbefundung (-4.853 engl. Pfund pro 100.000 Frauen ca. -0,07 € pro Frau). Empfehlen Doppelbefundung mit Konsensus als kosteneffektiv
Seradour et al., 1997, Frankreich	Analyse der realen additiven Kosten im Screening-Setting bei unilateraler Doppelbefundung (+15% Rate entdeckter Karzinome) <ul style="list-style-type: none"> ▪ direkte Kosten der Doppelbefundung ▪ indirekte Kosten der zusätzlichen Wiedereinbestellung (durch Fragebogen an einem Teil der Pat. erhoben und hochgerechnet): Bildgebung, falsch-positive Biopsien etc. 	Zusätzliche Kosten pro gescreenter Frau: 17,75 Francs (ca. 2,7 €) und 3.788 Francs pro zusätzlichem Karzinom (ca. 577 €) und 2.189 Francs pro falsch-positivem Befund Anteil Gesamtkosten am Screening: 5% Kosten bei dezentralem Setting insgesamt höher als bei zentralem Setting (Vergleich mit Anderson 1994) Erhöhte Kosten werden als gerechtfertigt bei höherer Rate entdeckter Karzinome eingeschätzt
Leivo et al.,	Kosten-Effektivitätsanalyse von	Durch Doppelbefundung wurden 290 statt 261

Autor, Jahr, Land	Inhalt	Ergebnis
1999, Finnland	unabhängiger und z.T. nicht unabhängiger Doppelbefundung , zum großen Teil unilateral, Vergleich zu Einzelbefundung 1990-1995 in 3 finnischen Screening-Zentren	<p>Karzinome (+11,1%) erkannt, darunter eine signifikant höhere Anzahl an in-situ-Karzinomen (40% versus 17%). Die Kosten pro entdecktem Karzinom waren 18,3 US\$ (ca. 13,4 €) Die Kosten pro durch Doppelbefundung entdecktem, zusätzlichem Karzinom waren 25,5 US\$ (rel. Anstieg 39%, ca. 18,7 €), bei Nicht-Einschluss von Ca-in-situ: 42,5 US\$ (ca. 31,1 €).</p> <p>Die zusätzlichen Kosten werden als nicht drastisch höher eingeschätzt, eingeräumt wird, dass die Kosten pro gewonnenem Lebensjahr sehr viel höher sein könnten. Bei Annahme, das Ca-in-situ, einer Over-Diagnosis entsprechen könnten, sind die Kosten deutlich erhöht.</p>
Tonita et al., 1999, Kanada	Retrospektive Analyse einer Doppelbefundung von 27.863 Mammographien durch speziell ausgebildete RTAs als zweiter Befunder	<p>Der PPV des Screening blieb gleich</p> <p>Speziell geschulte RTAs werden als kosteneffektiv zur Doppelbefundung erachtet.</p>
Ciatto et al., 2005a, Italien	<p>Retrospektive Studie von 1217 konsekutiven diskrepanten Doppelbefundungen mit Arbitrations – Vergleich der Kosten mit Unilateraler Doppelbefundung anhand Berechnung realer Ausgaben im italienischen Screening-Setting</p> <p>Hochrechnung von vorliegenden Befunden aus Screening-Situation pro 1.000 Frauen:</p>	<p>Aufgrund von Falsch-negativer Arbitration relative Einbusse der Rate entdeckter Karzinome um 1,7%</p> <p>Für jedes nicht-diagnostizierte Karzinom werden aufgrund der niedrigeren Recall-Rate 151 Recall und 21,23 € ‚eingespart‘, pro gescreenter Frau 1,72 €</p> <p>Halten Arbitration vor allem bei hoher Wiedereinbestellrate – z.B. bei neu eingeführtem Screening für kosteneffizient</p>
Ciatto et al., 2005b, Italien	Retrospektive Bewertung von nicht-unabhängiger unilateraler Doppelbefundung im Vergleich zu Einzelbefundung von 177.631 Mammographien 1998-2003. U.a. Kostenberechnung anhand realer Ausgaben im italienischen Screening-Setting.	<p>Kosten für Doppelbefundung waren 2,17 € zusätzlich pro untersuchter Frau und 11,168 € pro entdecktem Karzinom versus 11,585 € bei Einzelbefundung</p> <p>Unilaterale Doppelbefundung wird als kosteneffektiv bewertet</p>
Groenewoud et al., 2006 Niederlande	<p>Kosten-Effektivität von verschiedenen Befundungs- und Wiedereinbestellungsstrategien im Mammographie-Screening in den Niederlanden</p> <p>Fall-Kontroll Studie mit 26 Screening-Radiologen die 500 Mammographien befunden, 250 Kontrollen, 125 durch Screening entdeckte Karzinome und 125 mit Intervallkarzinomen</p> <p>1. Unabhängige Befundung</p> <p>2. Anwendung 3 verschiedener Befundungsstrategien der Doppelbefundung an den diskrepanten Fällen</p> <ul style="list-style-type: none"> - unilateral - Konsensus - Arbitration-Verfahren <p>Hochrechnung der Ergebnisse aus das niederländische Screening aufgrund mangelnder realer Zahlen mittels eines Microsimulationsmodells</p>	<p>Konkordanz von 2 Radiologen bei Kontrollen: 90,2% Bei Fällen: 75,2%</p> <p>Diskrepanz über alle Befundungen: 17,1%</p> <p>Unilaterales Doppelbefundung zeigte gegenüber Einzelbefundung die höchste Sensitivität</p> <p>Auf das Gesamt-Screening übertragen würde das 700 mehr gerettete Frauen bedeuten (12.000 Lebensjahre) bei einem relativen Anstieg der Wiedereinbestellung von 30%</p> <p>Die Kosten für unilaterale Doppelbefundung und für DR mit Konsensus (derzeitige Strategie in den Niederlanden) sind trotz der erhöhten Recall-Rate bei unilateralem Vorgehen etwa gleich: 2.168 € und 2.207 € pro gewonnenem Lebensjahr. Bei Einrechnung zusätzlicher invasiver Diagnostik steigen die Kosten auf 4.190 €</p>

Autor, Jahr, Land	Inhalt	Ergebnis
Lindfors et al., 6/2006, USA CAD	<p>Kosten-Effektivitäts-Analyse von Mammographie mit CAD-Befundung versus Mammographie mit Einzelbefundung versus Beobachtung (keine Mammographie). Markov-Modell. Sensitivitätsanalysen: Verschiedene Kosten für CAD, verschiedene Rate an Rate entdeckter Karzinome und Stadienverteilungen</p> <p>Berechnung der marginalen Kosten pro gerettetem Lebensjahr</p>	<p>Dies wird als akzeptabel angesehen</p> <p>Mammographie+CAD: 19,1 \$ (+19%. ca. 14,5 €, +2,3 €) Nur Mammographie: 16,0 \$ (ca. 12,2 €)</p> <p>Effectiveness von Screening mit CAD: +29% in dieser Simulation Kosten gesamt : +34%</p> <p>Nutzen-Effekt Abhängig von Rate entdeckter Karzinome und Stadienverteilung</p> <p>Erhöhung der Kosten für CAD erscheint in einem akzeptablen Rahmen</p>
Taylor et al., 2004 England	<p>HTA-Bericht zur Befundung mit CAD Fall-Kontroll Studie in 2 Settings mit 30 Radiologen, 5 klinischen Brustspezialisten und 15 RTAs</p> <ol style="list-style-type: none"> 1. Befundung von 180 Mammographien mit 40 durch Screening entdeckten Karzinomen und 20 Intervallkarzinomen. Befundung ohne und mit CAD-Reihenfolge randomisiert zugeteilt 2. Befundung von 120 Fällen von durch Radiologen übersehenen Befunden, die durch CAD angezeigt wurden 	<p>Studie 1: Sensitivität und Spezifität der Befundungen änderte sich nicht signifikant durch CAD (bei ausreichende Power für eine Signifikanzänderung von 0.1)</p> <p>Studie 2. Befundungen verbesserten sich durch CAD nicht signifikant</p> <p>Kosten: Hypothetische Berechnung nach einem Decision Analysis Modell von</p> <ol style="list-style-type: none"> 1. Kosten pro entdecktem Karzinom innerhalb von 12 Monaten + 23,2 engl. Pfund (ca. 37 €) 2. Kosten pro QUALY (nach 10 Jahren) +26,69 engl. Pfund (ca. 42,7 €) <p>CAD erhöht die Kosten bei keiner signifikanten Verbesserung in Sensitivität oder Spezifität</p>

7. Literatur

7.1 Evidenz Doppelbefundung

7.1.1 Systematischer Review zur Doppelbefundung

1. Dinnes J, Moss S, Melia J, Blanks R, Song F, Kleijnen J. Effectiveness and cost-effectiveness of double reading of mammograms in breast cancer screening: findings of a systematic review. *Breast* 2001;10(6):455-63

7.1.2 Eingeschlossene Studien zur Doppelbefundung

1. Anderson ED, Muir BB, Walsh JS, Kirkpatrick AE. The efficacy of double reading mammograms in breast screening. *Clin Radiol* 1994;49(4):248-51
2. Anttinen I, Pamilo M, Soiva M, Roiha M. Double reading of mammography screening films--one radiologist or two? *Clin Radiol* 1993;48(6):414-21
3. Ciatto S, Ambrogetti D, Bonardi R, Catarzi S, Risso G, Rosselli Del TM, Mantellini P. Second reading of screening mammograms increases cancer detection and recall rates. Results in the Florence screening programme. *J Med Screen* 2005;12(2):103-6
4. Ciatto S, Del Turco MR, Morrone D, Catarzi S, Ambrogetti D, Cariddi A, Zappa M. Independent double reading of screening mammograms. *J Med Screen* 1995;2(2):99-101
5. Deans HE, Everington D, Cordiner C. Scottish experience of double reading in the National Breast Screening Programme. *Breast* 1998;7(2):75-9
6. Duijm LE, Groenewoud JH, Hendriks JH, de Koning HJ. Independent double reading of screening mammograms in The Netherlands: effect of arbitration following reader disagreements. *Radiology* 2004;231(2):564-70
7. Kwek BH, Lau TN, Ng FC, Gao F. Non-consensual double reading in the Singapore Breast Screening Project: benefits and limitations. *Ann Acad Med Singapore* 2003;32(4):438-41
8. Liston JC, Dall BJ. Can the NHS Breast Screening Programme afford not to double read screening mammograms? *Clin Radiol* 2003;58(6):474-7
9. Pauli R, Hammond S, Cooke J, Ansell J. Comparison of radiographer/radiologist double film reading with single reading in breast cancer screening. *J Med Screen* 1996;3(1):18-22
10. Renaud R, Schaffer P, Gairard B, Dale G, Haehnel P, Kleitz C, Guldenfels C. [Principles and first results of the European program of breast cancer screening in the Bas-Rhin]. *Bull Acad Natl Med* 1991;175(1):129-45
11. Seradour B, Wait S, Jacquemier J, Dubuc M, Piana L. [Modalities of reading of detection mammographies of the programme in the Bouches-du-Rhone. Results and costs 1990-1995]. *J Radiol* 1997;78(1):49-54
12. Thurfjell E., Lernevall KA, Taube AA. Related Articles, Links Benefit of independent double reading in a population-based mammography screening program. *Radiology*. 1994 Apr;191(1):241-4
13. Warren RM, Duffy SW. Comparison of single reading with double reading of mammograms, and change in effectiveness with experience. *Br J Radiol* 1995;68(813):958-62
14. Williams SM, Doyle TCA, Chartres S. Impact of double reading of mammograms from the inception of a population-based breast cancer screening programme. *Breast* 1995;4(4):282-8

7.1.3 Ausgeschlossene Studien zur Doppelbefundung

1. Beam CA, Sullivan DC, Layde PM. Effect of human variability on independent double reading in screening mammography. *Acad Radiol* 1996;3(11):891-7
2. Blanks RG, Wallis MG, Moss SM. A comparison of cancer detection rates achieved by breast cancer screening programmes by number of readers, for one and two view mammography: results from the UK National Health Service breast screening programme. *J Med Screen* 1998;5(4):195-201
3. Ciatto S, Ambrogetti D, Risso G, Catarzi S, Morrone D, Mantellini P, Rosselli Del TM. The role of arbitration of discordant reports at double reading of screening mammograms. *J Med Screen* 2005;12(3):125-7
4. Denton ER, Field S. Just how valuable is double reporting in screening mammography? *Clin Radiol* 1997;52(6):466-8
5. Groenewoud JH, Otten JD, Fracheboud J, Draisma G, van Ineveld BM, Holland R, Verbeek AL, de Koning HJ. Cost-effectiveness of different reading and referral strategies in mammography screening in the Netherlands. *Breast Cancer Res Treat* 2006
6. Hukkinen K, Kivisaari L, Vehmas T. Impact of the number of readers on mammography interpretation. *Acta Radiol* 2006;47(7):655-9
7. Thurfjell E. Mammography screening. One versus two views and independent double reading. *Acta Radiol* 1994;35(4):345-50

7.1 Evidenz zu CAD-Systemen

7.2.1 Eingeschlossene Einzelstudien zu CAD-Systemen

1. Cupples TE, Cunningham JE, Reynolds JC. Impact of computer-aided detection in a regional screening mammography program. *AJR Am J Roentgenol* 2005;185(4):944-50
2. Dean JC, Ilvento CC. Improved cancer detection using computer-aided detection with diagnostic and screening mammography: prospective study of 104 cancers. *AJR Am J Roentgenol* 2006;187(1):20-8
3. Fenton JJ, Taplin SH, Carney PA, Abraham L, Sickles EA, D'Orsi C, Berns EA, Cutter G, Hendrick RE, Barlow WE, Elmore JG. Influence of computer-aided detection on performance of screening mammography. *N Engl J Med* 2007;356(14):1399-409
4. Freer TW, Ulissey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology* 2001;220(3):781-6
5. Gilbert FJ, Astley SM, McGee MA, Gillan MG, Boggis CR, Griffiths PM, Duffy SW. Single reading with computer-aided detection and double reading of screening mammograms in the United Kingdom National Breast Screening Program. *Radiology* 2006;241(1):47-53
6. Helvie MA, Hadjiiski L, Makariou E, Chan HP, Petrick N, Sahiner B, Lo SC, Freedman M, Adler D, Bailey J, Blane C, Hoff D, Hunt K, Joynt L, Klein K, Paramagul C, Patterson SK, Roubidoux MA. Sensitivity of noncommercial computer-aided detection system for mammographic breast cancer detection: pilot clinical trial. *Radiology* 2004;231(1):208-14
7. Khoo LA, Taylor P, Given-Wilson RM. Computer-aided detection in the United Kingdom National Breast Screening Programme: prospective study. *Radiology* 2005;237(2):444-9
8. Ko JM, Nicholas MJ, Mendel JB, Slanetz PJ. Prospective assessment of computer-aided detection in interpretation of screening mammography. *AJR Am J Roentgenol* 2006;187(6):1483-91
9. Morton MJ, Whaley DH, Brandt KR, Amrami KK. Screening mammograms: interpretation with computer-aided detection--prospective evaluation. *Radiology* 2006;239(2):375-83

7.2.2 Ausgeschlossene Studien zu CAD-Systemen

1. Baker JA, Rosen EL, Lo JY, Gimenez EI, Walsh R, Soo MS. Computer-aided detection (CAD) in screening mammography: sensitivity of commercial CAD systems for detecting architectural distortion. *AJR Am J Roentgenol* 2003;181(4):1083-8
2. Birdwell RL, Bandodkar P, Ikeda DM. Computer-aided detection with screening mammography in a university hospital setting. *Radiology* 2005;236(2):451-7
3. Bornefalk H. Implications of unchanged detection criteria with CAD as second reader of mammograms. *Med Phys* 2006;33(4):922-9
4. Brem RF, Baum J, Lechner M, Kaplan S, Souders S, Naul LG, Hoffmeister J. Improvement in sensitivity of screening mammography with computer-aided detection: a multiinstitutional trial. *AJR Am J Roentgenol* 2003;181(3):687-93
5. Ciatto S, Brancato B, Rosselli Del TM, Risso G, Catarzi S, Morrone D, Bricolo D, Zappa M. Comparison of standard reading and computer aided diagnosis (CAD) on a proficiency test of screening mammography. *Radiol Med (Torino)* 2003;106(1-2):59-65
6. Ciatto S, Rosselli Del TM, Burke P, Visioli C, Paci E, Zappa M. Comparison of standard and double reading and computer-aided detection (CAD) of interval cancers at prior negative screening mammograms: blind review. *Br J Cancer* 2003;89(9):1645-9
7. Ciatto S, Ambrogetti D, Collini G, Cruciani A, Ercolini E, Risso G, Rosselli Del TM. Computer-aided detection (CAD) of cancers detected on double reading by one reader only. *Breast* 2006;15(4):528-32
8. Destounis SV, DiNitto P, Logan-Young W, Bonaccio E, Zuley ML, Willison KM. Can computer-aided detection with double reading of screening mammograms help decrease the false-negative rate? Initial experience. *Radiology* 2004;232(2):578-84
9. Frigas A, Kapsimalakou S, Spyrou G, Koufopoulos K, Vassilaros S, Chatzimichael A, Mantas J, Ligomenides P. Evaluation of a breast cancer computer aided diagnosis system. *Stud Health Technol Inform* 2006;124:631-6
10. Hukkinen K, Pamilo M. Does computer-aided detection assist in the early detection of breast cancer? *Acta Radiol* 2005;46(2):135-9
11. Hukkinen K, Vehmas T, Pamilo M, Kivisaari L. Effect of computer-aided detection on mammographic performance: experimental study on readers with different levels of experience. *Acta Radiol* 2006;47(3):257-63
12. Karssemeijer N, Otten JD, Verbeek AL, Groenewoud JH, de Koning HJ, Hendriks JH, Holland R. Computer-aided detection versus independent double reading of masses on mammograms. *Radiology* 2003;227(1):192-200
13. Kegelmeyer WP, Jr., Pruneda JM, Bourland PD, Hillis A, Riggs MW, Nipper ML. Computer-aided mammographic screening for spiculated lesions. *Radiology* 1994;191(2):331-7
14. Lindfors KK, McGahan MC, Rosenquist CJ, Hurlock GS. Computer-aided detection of breast cancer: a cost-effectiveness study. *Radiology* 2006;239(3):710-7
15. Menna S, Di Virgilio MR, Burke P, Frigerio A, Boglione E, Ciccarelli G, Di FS, Garretti L. Diagnostic accuracy of commercial system for computer-assisted detection (CADx) as an adjunct to interpretation of mammograms. *Radiol Med (Torino)* 2005;110(4):334-40
16. Moberg K, Bjurstam N, Wilczek B, Rostgard L, Egge E, Muren C. Computed assisted detection of interval breast cancers. *Eur J Radiol* 2001;39(2):104-10
17. Pai VR, Gregory NE, Swinford AE, Rebner M. Ductal carcinoma in situ: computer-aided detection in screening mammography. *Radiology* 2006;241(3):689-94
18. Skaane P, Kshirsagar A, Stapleton S, Young K, Castellino RA. Effect of computer-aided detection on independent double reading of paired screen-film and full-field digital screening mammograms. *AJR Am J Roentgenol* 2007;188(2):377-84
19. Taplin SH, Rutter CM, Lehman CD. Testing the effect of computer-assisted detection on interpretive performance in screening mammography. *AJR Am J Roentgenol* 2006;187(6):1475-82

20. Taylor P, Given-Wilson R, Champness J, Potts HW, Johnston K. Assessing the impact of CAD on the sensitivity and specificity of film readers. *Clin Radiol* 2004;59(12):1099-105
21. Thurfjell E, Thurfjell MG, Egge E, Bjurstam N. Sensitivity and specificity of computer-assisted breast cancer detection in mammography screening. *Acta Radiol* 1998;39(4):384-8
22. Timp S, Karssemeijer N. Interval change analysis to improve computer aided detection in mammography. *Med Image Anal* 2006;10(1):82-95
23. Warren Burhenne LJ, Wood SA, D'Orsi CJ, Feig SA, Kopans DB, O'Shaughnessy KF, Sickles EA, Tabar L, Vyborny CJ, Castellino RA. Potential contribution of computer-aided detection to the sensitivity of screening mammography. *Radiology* 2000;215(2):554-62

7.2.3 Ausgeschlossene Reviews zu CAD-Systemen

1. Astley SM, Gilbert FJ. Computer-aided detection in mammography. *Clin Radiol* 2004;59(5):390-9
2. Balleyguier C, Boyer B, Athanasiou A, Vanel D, Sigal R. [Understanding CAD (computer-aided diagnosis) in mammography]. *J Radiol* 2005;86(1):29-35
3. Irwig L, Houssami N, van VC. New technologies in screening for breast cancer: a systematic review of their accuracy. *Br J Cancer* 2004;90(11):2118-22

7.3 Studien mit Kosten-Nutzen-Analysen

1. Anderson ED, Muir BB, Walsh JS, Kirkpatrick AE. The efficacy of double reading mammograms in breast screening. *Clin Radiol* 1994;49(4):248-51
2. Brown J, Bryan S, Warren R. Mammography screening: an incremental cost effectiveness analysis of double versus single reading of mammograms. *BMJ* 1996;312(7034):809-12
3. Ciatto S, Del Turco MR, Morrone D, Catarzi S, Ambrogetti D, Cariddi A, Zappa M. Independent double reading of screening mammograms. *J Med Screen* 1995;2(2):99-101
4. Ciatto S, Ambrogetti D, Bonardi R, Catarzi S, Risso G, Rosselli Del TM, Mantellini P. Second reading of screening mammograms increases cancer detection and recall rates. Results in the Florence screening programme. *J Med Screen* 2005b;12(2):103-6
5. Ciatto S, Ambrogetti D, Risso G, Catarzi S, Morrone D, Mantellini P, Rosselli Del TM. The role of arbitration of discordant reports at double reading of screening mammograms. *J Med Screen* 2005a;12(3):125-7
6. Groenewoud JH, Otten JD, Fracheboud J, Draisma G, van Ineveld BM, Holland R, Verbeek AL, de Koning HJ. Cost-effectiveness of different reading and referral strategies in mammography screening in the Netherlands. *Breast Cancer Res Treat* 2006;
7. Leivo T, Salminen T, Sintonen H, Tuominen R, Auerma K, Partanen K, Saari U, Hakama M, Heinonen OP. Incremental cost-effectiveness of double-reading mammograms. *Breast Cancer Res Treat* 1999;54(3):261-7
8. Lindfors KK, McGahan MC, Rosenquist CJ, Hurlock GS. Computer-aided detection of breast cancer: a cost-effectiveness study. *Radiology* 2006;239(3):710-7
9. Seradour B, Wait S, Jacquemier J, Dubuc M, Piana L. [Modalities of reading of detection mammographies of the programme in the Bouches-du-Rhone. Results and costs 1990-1995]. *J Radiol* 1997;78(1):49-54
10. Taylor P, Given-Wilson R, Champness J, Potts HW, Johnston K. Assessing the impact of CAD on the sensitivity and specificity of film readers. *Clin Radiol* 2004;59(12):1099-105
11. Tonita JM, Hillis JP, Lim CH. Medical radiologic technologist review: effects on a population-based breast cancer screening program. *Radiology* 1999;211(2):529-33

A Mammographie-Screening
bei Frauen 40-49 Jahre und >70 Jahre
Nutzen/Schaden-Bilanz

B Doppelbefundung von Screening-Mammographien
unter Berücksichtigung des
Entscheidungsverfahrens bei diskrepanten
Befunden und des Einsatzes von Computer-
Assistierten-Diagnose (CAD)-Systemen

C **Stellenwert der Mammasonographie im
Rahmen der Brustkrebsfrüherkennung**

D Indikationsstellung für die stereotaktische
Vakuumbiopsie im Vergleich zur stereotaktischen
Stanzbiopsie

E Nutzen der Exzision des Stichkanals nach minimal-
invasiver Biopsie der Brust zur Vermeidung von
Tumorzellverschleppung

F Her-2/neu-Bestimmung:
Vergleich der Methoden Immunhistochemie,
Fluoreszenz-in-situ-Hybridisierung und
Chromogene-in-situ-Hybridisierung

Inhalt

1. Hintergrund	84
2. Fragestellung	87
3. Methoden	87
3.1 Kriterien für den Einschluss und Ausschluss von Studien	87
3.1.1 Zielgruppe.....	88
3.1.2 Intervention und Referenzstandard	88
3.1.3 Studientypen.....	88
3.1.4 Erforderliche Angaben.....	89
3.1.5 Ein- und Ausschlusskriterien	89
3.2. Literaturrecherche.....	90
3.2.1 Datenbanken und Suchbegriffe	90
3.3 Zielgrößen und deren Definition	91
3.4 Bewertung der gefundenen Literatur	92
4. Ergebnisse	93
4.1 Stellenwert der ergänzenden Mammasonographie im Rahmen des Brustkrebs-Screenings	93
4.1.1 Ergebnisse der Literaturrecherche.....	93
4.1.2 Aufbereitete Evidenz	93
4.1.3 Evidenz von Einzelstudien	93
4.1.4 Diskussion.....	95
Literatur	97
4.2 Stellenwert der axillären Sonographie bei klinisch unauffälliger Axilla.....	97
4.2.1 Ergebnisse der Literaturrecherche.....	97
4.2.2 Aufbereitete Evidenz	97
4.2.3 Evidenz von Einzelstudien	98
4.2.4 Diskussion.....	99
4.3 Nutzen der Mammasonographie zur Karzinomdiagnostik bei auffälligem Brustbefund	100
4.3.1 Ergebnisse der Literaturrecherche.....	100
4.3.2 Aufbereitete Evidenz	101
4.3.3 Evidenz von Einzelstudien	101
4.3.4 Diskussion.....	102
4.4 Stellenwert der Mammasonographie bei Frauen unter 50 Jahren mit auffälligem Brustbefund	103
4.4.1 Ergebnisse der Literaturrecherche.....	103
4.4.2 Aufbereitete Evidenz	103
4.4.3 Evidenz von Einzelstudien	103
4.4.4 Diskussion.....	104
4.5 Anteil maligner Befunde bei mammographischer BI-RADS Klasse 3.....	105
4.5.1 Literaturrecherche.....	105

5. Evidenztabellen	106
5.1 Evidenz zur ergänzenden Mammasonographie beim Brustkrebs-Screening	106
5.1.1 Evidenz von Einzelstudien	106
5.1.2 Ausgeschlossene Einzelstudien	111
5.1.3 Ausgeschlossenen Reviews	111
5.2 Evidenz zur diagnostischen Sicherheit der Axillasonographie bei negativem klinischem Befund	112
5.2.1 Eingeschlossener Review	112
5.2.2 Evidenz von Einzelstudien	113
5.2.3 Ausgeschlossene Studien	119
5.3 Mammasonographie bei auffälligem Brustbefund	122
5.3.1 Evidenz von Einzelstudien	122
5.3.2 Ausgeschlossene Studien	131
5.4 Stellenwert der Mammasonographie Frauen unter 50 Jahren mit auffälligem Brustbefund	133
5.4.1 Evidenz von Einzelstudien	133
6. Literatur	140

Tabellenverzeichnis

Tabelle 1	Evidenz von Einzelstudien zur ergänzenden Mammasonographie beim Brustkrebs-Screening	106
Tabelle 2	Ausgeschlossene Einzelstudien zur ergänzenden Mammasonographie beim Brustkrebs-Screening	111
Tabelle 3	Ausgeschlossene Reviews zur ergänzenden Mammasonographie beim Brustkrebs-Screening	111
Tabelle 4	Systematischer Review zur Axillasonographie	112
Tabelle 5	Evidenz von Einzelstudien zur Axillasonographie	113
Tabelle 6	Ausgeschlossene Einzelstudien zur Axillasonographie	119
Tabelle 7	Ausgeschlossene Reviews zur Axillasonographie	121
Tabelle 8	Evidenz von Einzelstudien zur Mammasonographie bei auffälligem Brustbefund	122
Tabelle 9	Ausgeschlossene Einzelstudien zu Mammasonographie bei auffälligem Befund	131
Tabelle 10	Ausgeschlossene Reviews zu Mammasonographie bei auffälligem Brustbefund	132
Tabelle 11	Evidenz von Einzelstudien zur Mammasonographie bei Frauen unter 50 Jahren mit auffälligem Brustbefund	133

1. Hintergrund

Die Mammasonographie wurde bis Anfang der 90er Jahre vorwiegend zur Differenzierung von Zysten und soliden Befunden sowie zur bildgebenden Steuerung minimal-invasiver Interventionen eingesetzt [1]. Für die Detektion von Karzinomen wurde im Vergleich zur Mammographie eine schlechte Sensitivität mit hohen Raten falsch-positiver Befunde berichtet [1-3].

Durch die Weiterentwicklung der Technik, vor allem aufgrund der Einführung der kontrastverstärkten und hochauflösenden Sonographie mit Schallkopffrequenzen größer fünf Megahertz [4], hat sich das diagnostische Potenzial der Mammasonographie seit Anfang der 90er Jahre stark verbessert:

- Stavros et al. publizierten 1995 eine prospektive Studie (n= 622) zur Klassifizierung palpabler und nicht palpabler sonographisch solider Brustbefunde unter Verwendung von Hochfrequenz-Schallköpfen von 7,5-10 Mhz. Bei einer Karzinomprävalenz von 17% erzielten sie eine Sensitivität von 98,4% und einen negativen prädiktiven Wert von 99,5% für die Erkennung maligner Befunde [5]. Ein wichtiger Aspekt der Arbeit war die genaue Definition und Re-Evaluierung von sonomorphologischen Kriterien zu Benignität und Malignität.
- In Studien anderer Arbeitsgruppen erfolgte ebenfalls eine Evaluierung sonomorphologischer Kriterien bei soliden Befunden. Skaane et al., 1998, erzielten in einer prospektiven Studie (n= 362) eine Sensitivität von 99,5% und einen negativen prädiktiven Wert von 98% [6]. Rahbar et al., 1999, gaben in einer retrospektiven Auswertung in Bezug auf einzelne Kriterien eine maximale Sensitivität bis zu 94% an [7].

Die Mammasonographie wurde in diesen Studien nicht verblindet zur Mammographie durchgeführt.

Aufgrund der verbesserten Sensitivität in Bezug auf eine Differenzierung zwischen malignen und benignen Befunden hat sich die Mammasonographie als ergänzendes bildgebendes Verfahren zur Mammographie in den letzten Jahren breit etabliert [8].

Um einheitliche Bewertungskriterien zu gewährleisten, entwickelte das American College of Radiology (ACR) 2003 eine BI-RADS (Breast Imaging Reporting And Data System) – Klassifikation für Mammasonographie analog der BI-RADS – Klassifikation für Mammographien [9]. Die mammasonographische BI-RADS-Klassifikation beinhaltet fünf Befundkategorien: 1. unauffällig, 2. benigne, 3. wahrscheinlich benigne, 4. suspekt und 5. wahrscheinlich maligne. Eine zusätzliche Kategorie 0 wurde für Befunde vergeben, bei denen aufgrund eingeschränkter Beurteilung eine weitere Bildgebung erforderlich ist. Für jede Befundkategorie sind detaillierte sonographische Kriterien beschrieben. Diese Klassifikation wurde 2006 von der Deutschen Gesellschaft für Ultraschall in der Medizin (DEGUM) adaptiert übernommen [10].

Für die Früherkennung von Brustkrebs mit dem Ziel einer Prognoseverbesserung für die betroffenen Frauen müssen vor allem kleine Befunde von möglichst kleiner 1 cm, maximal jedoch 2 cm Größe identifiziert werden können (T1 Tumoren). Die wichtigsten Prognosefaktoren im Hinblick auf eine Heilung sind die Größe des Tumors und der axilläre Lymphknotenbefall. Die Wahrscheinlichkeit des Lymphknotenbefalls erhöht sich mit der Größe des Tumolvolumens [11].

Mammographisch werden etwa 35-45% nicht palpabler suspekter Befunde durch Mikrokalzifikationen entdeckt [12]. Solche Mikrokalzifikationen können auch mit modernen Mammasonographiegeräten nicht zuverlässig dargestellt werden [13,14].

Vor diesem Hintergrund nennt die Stufe-3-Leitlinie Brustkrebsfrüherkennung in Deutschland von 2003 als Schlüsselempfehlung zur Indikation für die Mammasonographie die Abklärung unklarer Befunde im Sinne einer Zusatzuntersuchung und wertet sie als ungeeignet als primäre Methode der Früherkennung von Mammakarzinomen [8].

Für die Aktualisierung der Leitlinie forderte die Expertengruppe Mammasonographie der Konzentrierten Aktion Brustkrebsfrüherkennung für folgende Themen aktuelle Evidenzanalysen hinsichtlich der Indikationsstellung zur Mammasonographie an:

1. Nutzen der zusätzlichen Mammasonographie bei asymptomatischen Frauen im Rahmen des populationsbezogenen Brustkrebs-Screenings.
2. Diagnostische Sicherheit der Axillasonographie als bildgebendes Verfahren bei der Brustkrebsfrüherkennung bei negativem klinischem Befund
3. Nutzen der Mammasonographie zur Karzinomdiagnostik bei auffälligem Brustbefund.
4. Stellenwert der Mammasonographie bei Frauen unter 50 Jahren mit auffälligem Brustbefund.
5. Anteil der Malignität bei mammographischen BI-RADS- Klasse 3 – Befunden.

Im Folgenden werden die Fragestellungen erläutert:

1. Nutzen der zusätzlichen Mammasonographie bei asymptomatischen Frauen im Rahmen des populationsbezogenen Brustkrebs-Screenings

Beim Brustkrebs-Screening durch Mammographie werden nicht alle Karzinome erkannt. Das ergaben retrospektive Analysen des Bildmaterials von Karzinomen, die im Screening-Intervall oder in einer weiteren Screening-Runde diagnostiziert wurden [15]. Bei nicht erkannten Karzinomen handelt es sich in etwa zwei Drittel um Architekturstörungen und Herdbefunde und in etwa einem Drittel um Mikrokalzifikationen [16]. Die Auswirkungen des zusätzlichen Einsatzes der Mammasonographie in Bezug auf Sensitivität und Spezifität des Screenings sollten bestimmt werden.

2. Diagnostische Sicherheit der Axillasonographie bei negativem palpablem Befund als bildgebendes Verfahren bei der Brustkrebsfrüherkennung

Die klinische Beurteilung der axillären Region bietet in Bezug auf die Erkennung von Lymphknotenmetastasen eine schlechte Sensitivität [17], ebenso die Mammographie [18]. Da der Lymphknotenstatus neben der Größe des Tumors wichtigster Prognosefaktor ist, wird er derzeit bei allen Patientinnen mit invasivem Mammakarzinom operativ durch eine klassische Axilladissektion oder durch die Entfernung des so genannten Sentinel-Nodes (Wächterlymphknoten, die erste Lymphknotenstation des Tumorabflusses) bestimmt [19]. Im Rahmen der Brustkrebsfrüherkennung interessierte primär die diagnostische Sicherheit einer unauffälligen Axillasonographie bei kleinen Tumoren bis zwei Zentimeter (T1) und klinisch negativer Axilla. Die Prävalenz von Lymphknotenmetastasen bei Tumoren bis zwei Zentimeter ist gering, sie liegt bei ca. 30% [11]. Mehr als zwei Drittel der Patientinnen werden demzufolge operiert, obwohl sie einen unauffälligen Befund haben.

3. Nutzen der Mammasonographie für die Karzinomdiagnostik bei auffälligem Brustbefund

Die Diagnosestellung Brustkrebs bei bereits palpablen Befunden kann generell zur Reduktion der Brustkrebsmortalität keinen wesentlichen Beitrag leisten [20]. Im Rahmen einer Evidenzanalyse sollte jedoch aufgezeigt werden, welche Sensitivität und Spezifität die Mammasonographie für die klinisch auffälligen Befunde im Unterschied zu der Mammographie liefert. Es interessierte darüber hinaus die Größe der sonographisch diagnostizierten Karzinome.

4. Stellenwert der Mammasonographie bei Frauen unter 50 Jahren mit auffälligem Brustbefund

Die Stufe-3-Leitlinie Brustkrebsfrüherkennung in Deutschland 2003 nennt in der Auflistung der Einzelindikationen die Untersuchung junger Frauen mit Tastbefund als einzige Indikation für den primären Einsatz der Mammasonographie [8]. Begründet werden kann diese Altersangabe mit der geringen Inzidenz von Brustkrebs in der Altersgruppe unter 40 Jahren [11] und dem gleichzeitig erhöhten Anteil von Frauen mit mammographisch dichter Brust und dadurch erschwerter mammographischer Beurteilbarkeit [21]. Die Strahlenbelastung durch eine Mammographie erscheint somit bei diesem Kollektiv nicht gerechtfertigt. Die Empfehlung sollte einer Evidenzüberprüfung unterzogen werden.

5. Anteil der Malignität bei mammographischen BI-RADS- Klasse 3 – Befunden

In der Stufe-3-Leitlinie Brustkrebsfrüherkennung in Deutschland 2003 wird die zusätzliche Mammasonographie bei mammographischen Befunden der BI-RADS Klasse 4 und 5 routinemäßig zur weiteren Befunddifferenzierung empfohlen [8]. In diesen Befundklassen findet sich ein hoher Anteil an Karzinomen, die Befunde sollen zwingend histologisch abgeklärt werden [9]. Bei den Befunden der BI-RADS-Klasse 3, die als wahrscheinlich gutartig eingestuft werden, wird eine Malignitätsrate von maximal 2% angenommen. Diese Befunde sollen keiner histologischen Abklärung, sondern lediglich einer mammographischen Kontrolle zugeführt werden [9]. Für diese Befunde wurde bisher keine Empfehlung einer Routinemammasonographie ausgesprochen. Die in Studien tatsächlich berichtete Anzahl an malignen Befunden bei BI-RADS 3 sollte als Grundlage für die Beibehaltung oder Änderung der bisherigen Empfehlung erhoben werden.

Literatur

1. Jackson VP. The current role of ultrasonography in breast imaging. *Radiol Clin North Am* 1995;33(6):1161-70
2. Sickles EA, Filly RA, Callen PW. Breast Cancer Detection with sonography and mammography: comparison using state-of-the-art equipment. *AJR* 1983;140:843-5
3. Bassett LW, Kimme-Smith C. Breast Sonography. *AJR* 1991;156:449-456
4. ESAOTE. A history of People and Technology. Strategic Corporate Marketing. Genoa: Esaote Spa; 2005
5. Stavros AT, Thickman D, Rapp CL, Dennis MA, Parker SH, Sisney GA. Solid Breast Nodules: Use of Sonography to distinguish between Benign and Malign Lesions. *Radiology* 1995; 196(1):123-34
6. Skaane P, Engedal K. Analysis of sonographic features in the differentiation of fibroadenomas and invasive ductal carcinoma. *AJR* 1998;170:109-14
7. Rahbar G, Sie AC, Hansen GC, Prince JS, Melany ML, Reynolds HE, Jackson VP, Syre JW, Bassett LW. Benign versus malignant solid breast masses: US differentiation. *Radiology* 1999;213(3):889-94
8. Schulz KD, Albert US (Hrsg.). Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland. München: Zuckschwerdt; 2003
9. American College of Radiology (ACR). Breast Imaging Reporting And Data System (BI-RADS™) 4th Edition. Reston (VA): American College of Radiology; 2003
10. Madjar H, Ohlinger R, Mundinger A, Watermann D, Frenz JP, Bader W, Schulz-Wendtland R, Degenhardt F. BI-RADS-analogue DEGUM criteria for findings in breast ultrasound – consensus of the DEGUM Committee on Breast Ultrasound. *Ultraschall Med* 2006;27(4):374-379
11. Sauer H. (Hrsg.). Manual Mammakarzinome: Empfehlungen zur Diagnostik Therapie und Nachsorge. 10. Auflage. München: Zuckschwerdt; 2005
12. Sickles EA. Mammographic detectability of breast microcalcifications. *AJR* 1982;139:913-8
13. Cheung YC, Wan YL, Chen SC, Lui KW, Ng SH, Yeow KM, Lee KF, Hsueh S. Sonographic evaluation of mammographically detected microcalcifications without a mass prior to stereotactic core needle biopsy. *J Clin Ultrasound* 2002;30(6):323-31
14. Soo MS, Baker JA, Rosen EL. Sonographic detection and sonographically guided biopsy of breast microcalcifications. *AJR* 2003;180(4):941-8
15. Boyr B, Hauret L, Bellaiche R, Graf C, Bourcier B, Fichet G. Cancers rétrospectivement visibles: revue de la littérature. *J Radiol* 2004;85(12 Pt 2):2071-8
16. Ikeda DM, Anderson I., Wattsgard C, Janzon L, Linell F. Interval Carcinomas in the Malmo Screening Trial: radiographic appearance and Prognostic considerations. *AJR* 1992;159(2):287-94
17. Lannig C, Hoffmann J, Galatius H, Engel U. Assessment of clinical palpation of the axilla as a criterion for performing the sentinel node procedure in breast cancer. *Eur J Surg Oncol.* 2007; 33 (3):281-4
18. Walsh R, Kornguth PJ, Soo MS, Bentley R, DeLong DM. Axillary lymph nodes: mammographic, pathologic, and clinical correlation. *AJR* 1997;168(1):33-8

19. Deutsche Krebsgesellschaft (Hrsg.). Diagnostik, Therapie und Nachsorge des Mammakarzinoms der Frau. München: Zuckschwerdt; 2004
20. Kusters JP, Gotzsche PC.. Regular self-examination or clinical examination for early detection of breastcancer. Cochrane Database Syst Rev 2003;(2):CD003373
21. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, Geller BM, Abraham LA, Taplin SH, Dignan M, Cutter G, Ballard-Barbash R. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Ann Intern Med 2003;138(3):168-75

2. Fragestellung

Die Bearbeitung der genannten Themenbereiche erfolgte auf der Grundlage folgender Schlüsselfragen, die von der Methodikerin in Abstimmung mit der Expertengruppe Mammasonographie der konzertierten Aktion Brustkrebsfrüherkennung generiert wurden:

1. Kann eine zusätzliche Mammasonographie beim Brustkrebs-Screening (kein familiär bedingtes Hochrisiko-Kollektiv) die Karzinomentdeckung verbessern?
Welcher potentielle Schaden entsteht für die untersuchten, asymptomatischen Frauen?
2. Welche diagnostische Sicherheit weist die Sonographie im Hinblick auf die Beurteilung axillärer Lymphknoten bei klinisch negativer Axilla auf?
3. Welchen Nutzen für die Diagnostik von Karzinomen hat die Mammasonographie bei auffälligem Brustbefund?
4. Welchen Stellenwert hat die Mammasonographie bei der Abklärung auffälliger Brustbefunde bei Frauen unter 50 Jahren?
5. Welchen Anteil an malignen histologischen Befunden weisen mit BI-RADS Klasse 3-klassifizierte Mammographien auf? ¹

3. Methoden

3.1 Kriterien für den Einschluss und Ausschluss von Studien

Ziel war es, Studien zu identifizieren, die zur Frage des Nutzens (positive Effekte einer Intervention) der Mammasonographie im Hinblick auf die spezifischen Fragestellungen (siehe 2) unter Berücksichtigung etwaiger Schäden (negative Effekte einer Intervention) verlässliche Aussagen machen.

Im Folgenden sind die Kriterien für den Einschluss bzw. Ausschluss von Studien dargelegt. Die Kriterien wurden prospektiv von der Methodikerin in Abstimmung mit der Expertengruppe Mammasonographie vor der systematischen Literaturrecherche festgelegt.

¹ **Anmerkung:** Für diese Frage wurde vom ÄZQ nur die Primärrecherche der Literatur und das Screening der Abstracts durchgeführt, diese Ergebnisse sind unter 4.5. bzw. unter 6. Literatur aufgeführt. Die Studienbewertung erfolgte durch die Expertengruppe Mammasonographie. Die Ergebnisse sind nicht im vorliegenden Bericht enthalten.

3.1.1 Zielgruppe

Die Zielgruppe für die vorliegenden Evidenzanalysen waren Frauen ohne familiäres Risiko für Brustkrebs. Frauen mit familiär bedingter Hochrisikosituation für Brustkrebs stellen eine spezielle Gruppe dar, deren Ergebnisse in Bezug auf die Brustkrebsfrüherkennung aufgrund der erhöhten Karzinominzidenz nicht auf die Gesamtpopulation übertragbar sind [22;23].

Für die einzelnen Fragestellungen wurden die einzuschließenden Studienpopulationen spezifiziert.

zu 1) Asymptomatische Frauen mit Screening-Mammographie.

Unter Screening-Mammographie werden hier Mammographien an asymptomatischen Patientinnen verstanden. Eine Einschränkung der Bewertung auf eingeladene Frauen eines organisierten Screening-Programms erfolgte nicht.

zu 2) Patientinnen mit primärem, nicht metastasiertem Brustkrebs und klinisch negativer Axilla ohne neoadjuvante Chemotherapie, d.h. der axilläre Befund entspricht dem Status bei Diagnosestellung.

zu 3) Frauen mit entweder selbst oder durch klinische Untersuchung erhobenem auffälligen Brustbefund.

Studien mit Studienpopulationen, in der alle Studienteilnehmerinnen unter 50 Jahre alt waren, wurden unter 4) erfasst.

zu 4) Mammasonographie bei Frauen unter 50 Jahren mit palpablem Brustbefund – siehe 3).

zu 5) Anteil maligner mammographischer BI-RADS-Klasse 3-Befunde.

Alle Studien, die eine Bewertung von mammographischen BI-RADS- Klasse 3-Befunden beinhalteten, wurden im Volltext an die Expertengruppe Mammasonographie weitergegeben und nicht im ÄZQ analysiert. Eine Einschränkung in Bezug auf das Alter der untersuchten Patientinnen erfolgte nicht.

3.1.2 Intervention und Referenzstandard

Die Fragestellungen beziehen sich auf die Intervention Mamma- und Axillasonographie. Die eingesetzten Sonographiegeräte mussten in Bezug auf die Frequenz des Schallkopfes dem geltenden Standard entsprechen. Bei Verwendung eines Schallkopfes mit nur einer Frequenz mußte diese mehr als fünf Megahertz betragen, bei Verwendung eines Mehrfrequenzschallkopfes mußte die Frequenz fünf Megahertz und mehr betragen. Der Stellenwert der Intervention 'Dopplersonographie der Mamma oder der Axilla' war nicht Teil der Fragestellung. Die histologische Sicherung oder die adäquate Nachbeobachtung (siehe auch 3.3 Zielgrößen und ihre Definition) waren der Referenzstandard. Die Vergleichsbehandlung waren die Mammographie oder die Palpation (siehe jeweilige Frage).

3.1.3 Studientypen

Zur Beantwortung der vorliegenden Fragestellungen außer Frage 5 liefern randomisierte kontrollierte Studien (RCT) und ihre aggregierte Auswertung in Form systematischer Übersichtsarbeiten und Meta-Analysen die zuverlässigsten Ergebnisse [24]. Die Ergebnisse solcher Studien sind mit dem geringsten Risiko einer systematischen Verzerrung behaftet. Da eine orientierende Vorrecherche zeigte, dass zum Recherchezeitpunkt zu den vorliegenden Fragen nur einzelne, evtl. auch keine RCTs existierten, wurden zunächst konsekutive, prospektive Kohortenstudien berücksichtigt, danach nicht-konsequente und/oder retrospektive Kohortenstudien.

Nur bei sehr schlechter Evidenzlage aufgrund von Kohortenstudien wurden nachrangig auch Fall-Kontrollstudien berücksichtigt, wenn die Studiengruppen so beschrieben waren, dass von einer Vergleichbarkeit der Fall- und der Kontrollgruppe ausgegangen werden konnte. Fallberichte und Expertenmeinungen wurden von der Bewertung ausgeschlossen (siehe 3.4 und Anhang 1: Oxford-Klassifikation).

3.1.4 Erforderliche Angaben

Folgende zum Studieneinschluss erforderlichen Angaben wurden vorab im Hinblick auf die Auswertbarkeit und Vergleichbarkeit der Studien festgelegt:

a) für alle Fragestellungen erforderlich

- Größe der Studienpopulation (nicht nur die Anzahl der erhobenen Befunde),
- verwendete Schallkopf-Frequenz (Mhz),
- ausgewiesene Anzahl von Befunden für die Intervention Mammasonographie oder Axillasonographie allein

Das zunächst festgelegte Kriterium der Angabe von Größe oder Tumorstadium für die histologischen Befunde konnte für Frage 3 und 4 nicht aufrechterhalten werden, da die Angabe in 7 von 11 Studien (Frage 3) bzw. bei allen identifizierten Studien (Frage 4) fehlten. Dieses Kriterium wurde deshalb nur auf die Fragen 1 und 2 angewendet.

Ein zweites vorab festgelegtes Kriterium: ‚Anzahl und Erfahrung der Befunder‘ wurde ebenso zurückgezogen, da die Angaben dazu, wenn sie vorhanden waren, zu wenig quantifizierbar waren. Oft wurde nur ‚erfahren‘ oder ‚unerfahren‘ angegeben. Deshalb wird die Angabe zwar in den Tabellen aufgeführt, wurde aber nicht als zwingend vorausgesetzt.

b) zu den einzelnen Fragestellungen erforderliche Angaben:

zu 1) Brustdichte nach BI-RADS

Die mammographische Brustdichte kann nach einer BI-RADS-Klassifikation des ACR eingeteilt werden. Man unterscheidet 4 Befundklassen: 1. die Brust besteht nahezu vollständig aus Fettgewebe, 2. es finden sich verstreute fibroglanduläre Verdichtungen 3. das Brustgewebe ist von heterogener Dichte, 4. extrem dichtes Brustgewebe [9]. Die mammographische Sensitivität korreliert invers zur Brustdichte. Am sensitivsten ist die Mammographie zur Beurteilung von Befunden der Brustdichte BI-RADS-Klasse 1 [21].

zu 2)

Zusätzlich zu den unter a) genannten Kriterien sollte angegeben sein, ob der klinische Axillabefund negativ war oder ob ein palpabler axillärer Befund vorlag.

Darüber hinaus war die Angabe der sonographischen Kriterien für Malignität erforderlich.

Zunächst wurde als weiteres Kriterium die Angabe des Anteils dargestellter Lymphknoten festgelegt. Da jedoch nur einige der Studien dieses Kriterium erfüllten, wurde es bei einer insgesamt kleinen Anzahl zur Verfügung stehender Studien zurückgezogen.

3.1.5 Ein- und Ausschlusskriterien

Es galten folgende **Einschlusskriterien**:

1. Studie/ Review bearbeitet die Fragestellung
2. adäquater Studientyp
3. adäquate Studienpopulation
4. Intervention erfüllt die technische Anforderung
5. erforderliche Angaben sind genannt

Alle Einschlusskriterien mussten erfüllt sein.

Es galten folgende **Ausschlusskriterien**:

1. Doppelpublikationen
2. nicht im festgelegten Recherchezeitraum publiziert
3. methodisch: Fallberichte und Expertenmeinungen sowie qualitativ minderwertige Fall-Kontroll-Studien

Keines der Ausschlusskriterien durfte vorliegen.

3.2. Literaturrecherche

Die systematische Literaturrecherche im ÄZQ erfolgte für die Fragen 1 bis 4 zunächst für den Zeitraum ab Januar 2000. Für Frage eins und zwei wurden Publikationen bis März 2007 und für Frage drei bis fünf bis April 2007 im ÄZQ (jeweiliges Rechercheschlussdatum siehe unter den entsprechenden Ergebnissen der Literaturrecherchen). Aufgrund der geringen Zahl und der Heterogenität der Studien wurde der Suchzeitraum in einem zweiten Schritt für Frage 3 bis Januar 1999 und für Frage 4 bis Januar 1998 ausgeweitet. Die Recherche für Frage 5 erfolgte ohne zeitliche Beschränkung.

Die Abstracts der jeweiligen Literatursuche wurden von der Methodikerin auf inhaltliche Relevanz geprüft, nicht relevante Abstracts wurden ausgeschlossen. Inhaltlich als relevant beurteilte Abstracts wurden als Volltexte hinsichtlich der Ein- und Ausschlusskriterien geprüft. Der Ausschluss von Volltexten wurde jeweils begründet.

3.2.1 Datenbanken und Suchbegriffe

In folgenden Datenbanken wurde nach Vorgabe des Auftraggebers eine systematische Literaturrecherche durchgeführt:

1. PubMed (Internetportal der National Library of Medicine)
(<http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed>)
2. Database of Abstracts of Reviews of Effects (DARE), des Centre for Reviews and Dissemination (CRD)
(<http://www.york.ac.uk/inst/crd/crddatabases.htm#DARE>)
3. Cochrane-Datenbanken 'Cochrane Reviews' und 'Clinical Trials'
(<http://www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/HOME>).

Folgende Suchbegriffe wurden verwendet:

a) Für PubMed:

zu 1) screening AND (mammography OR mammogram) AND breastneoplasms
AND (ultrasonography, mammary OR breast echography)

zu 2) sonography OR ultrasound or ultrasonograph* OR echograph*) AND (axilla OR
axillary) AND breast AND (predictive OR value OR specific* OR accuracy OR
detection OR metasta* OR evaluat*)

zu 3) und 4)

breast AND (mass or masses or thickening OR lump OR symptom* OR
palpable) AND (sonography mammary OR ultrasound OR echography OR
sonography) AND (predictive value OR sensitivity OR specificity OR
accuracy OR false negative OR false positive)

zu 5) bi-rads AND (cancer OR malign* OR breast neoplasms)

b) Für Cochrane und DARE

zu 1) mammography screening AND ultrasound

zu 2) ultrasound AND axilla / ultrasound AND sentinel AND breast

zu 3) und 4)

ultrasound AND diagnosis AND breast AND cancer

zu 5) bi-rads AND breast AND cancer

3.3 Zielgrößen und deren Definition

Entsprechend der oben genannten Fragestellungen wurden durch die Methodikerin in Abstimmung mit dem Auftraggeber folgende Zielgrößen festgelegt:

zu 1)

Relevanter Endpunkt für die Beurteilung des Nutzens der zusätzlichen Mammasonographie ist die Rate an Brustkrebsmortalität. Die Kenntnis dieses Wertes ermöglicht die Beurteilung, ob durch eine zusätzlich durchgeführte Sonographie die Sensitivität für die Diagnose früher Formen des Brustkrebses so verbessert wird, dass die Mortalitätsrate an Brustkrebs gesenkt wird. Da keine der identifizierten Studien zu diesen Endpunkten Stellung nimmt, wurden als Surrogatparameter die Bewertungskriterien diagnostischer Tests (s.u.) und die zusätzliche relative sowie absolute Rate an entdeckten Karzinomen im Vergleich zur Mammographie festgelegt.

Der negative prädiktive Wert, die Sensitivität und Spezifität der Mammasonographie wurden nicht berechnet, wenn nicht eine Nachbeobachtung von mindestens dem Zeitraum eines Screening-Intervalls vorlag, und somit eine Ermittlung falsch-negativer Befunde möglich war.

Der Schaden für die untersuchten Frauen wurde anhand der Anzahl von aufgrund der Mammasonographie durchgeführten Biopsien und des positiven prädiktiven Werts dieser Biopsien im Hinblick auf die Malignität des histologischen Befundes quantifiziert, d.h. es wurde untersucht, wie viel unnötige Biopsien aufgrund eines falsch-positiven Befunds durchgeführt wurden.

Für eine umfassende Bewertung des Schadens durch die zusätzliche mammasonographische Bildgebung müsste als Endpunkt die zusätzliche Rate an Brustkrebs-Überdiagnostik ermittelt werden. Überdiagnostik bedeutet die Diagnose nicht klinisch relevanter Mammakarzinome [25]. Zu diesem Endpunkt nahm keine der Studien Stellung.

zu allen Fragen (1-4):

Es gelten die Beurteilungskriterien für diagnostische Tests, wie sie im Folgenden definiert sind [25]. Zu beachten ist der Einfluss der Prävalenz der untersuchten Erkrankung in einem Kollektiv bzw. einer Studienpopulation in Bezug auf den positiven und negativen prädiktiven Wert.

⇒ **positiver prädiktiver Wert:**

$$\frac{\text{Anzahl richtig positiver Befunde}}{\text{Anzahl aller positiven Befunde}}$$

⇒ **negativer prädiktiver Wert:**

$$\frac{\text{Anzahl richtig negativer Befunde}}{\text{Anzahl richtig negativer Befunde} + \text{Anzahl falsch negativer Befunde}}$$

⇒ **Sensitivität:**

$$\frac{\text{Anzahl richtig positiver Befunde}}{\text{Anzahl richtig positiver Befunde} + \text{Anzahl falsch negativer Befunde}}$$

⇒ **Spezifität:**

$$\frac{\text{Anzahl richtig negativer Befunde}}{\text{Anzahl richtig negativer Befunde} + \text{Anzahl falsch positiver Befunde}}$$

3.4 Bewertung der gefundenen Literatur

Die identifizierten Studien wurden nach Vorgabe des Auftraggebers mit der Oxfordklassifikation für Diagnosestudien bewertet [<http://www.cebm.net/index.aspx?o=1025>; Levels of Evidence und siehe Anhang 1]. Nach dieser Klassifikation wird Evidenz aus Diagnosestudien anhand von 5 Evidenzgraden bewertet.

Die Klassifikation für Diagnosestudien sieht im Unterschied zu der Oxford-Klassifikation für therapeutische Studien die Beurteilung von randomisierten kontrollierten Studien nicht vor. Evidenzstufe 1 wird für Validierungskohortenstudien (1b) oder systematische Reviews von Validierungskohortenstudien (1a) vergeben.

Die Klassifikation für Diagnosestudien berücksichtigt bei der Bewertung – im Unterschied zur Bewertung von Differentialdiagnosestudien – nicht, ob diese prospektiv oder retrospektiv erstellt wurden. Die prospektive Erhebung von Daten wird jedoch allgemein als wichtiges Gütekriterium von Studien gesehen [26]. Beispielhaft sei für die vorliegende Fragestellung die eingeschränkte retrospektive Beurteilbarkeit von sonographischen Befunden genannt.

Für die vorliegende Auswertung wurden deshalb prospektiv erhobene Studien vor den retrospektiven Studien dargestellt und bei der Bewertung wurde geprüft, ob in prospektiven Studien andere Ergebnisse erzielt wurden als in retrospektiven.

Nach der Oxford-Klassifikation zählt als Qualitätsmerkmal für eine Validierungskohortenstudie (Evidenzstufe 1) oder eine explorative Kohortenstudie (Evidenzstufe 2) neben einem unabhängigen, guten Referenzstandard auch, ob die eingeschlossenen Patientinnen konsekutiv rekrutiert wurden und ob verblindet diagnostiziert wurde. Bei einer Fragestellung, (Frage 1 und 2) die als Voraussetzung für die Intervention (Mamma- bzw. Axillasonographie) unauffällige vorausgehende, diagnostische Befunde (unauffällige Mammographie oder Palpation) voraussetzte, erschien die Verblindung kein zwingendes Merkmal und wurde deshalb nicht in die Bewertung miteinbezogen.

Durch die Oxford-Klassifikation wird die Publikationsqualität einer Studie nur zum Teil abgebildet. Wichtige Angaben für Studien zu diagnostischen Tests in der Onkologie sind in der Regel die Altersverteilung des Studienkollektivs, die quantifizierbare Angabe der Erfahrung der Befunder sowie die Tumorstadienverteilung bzw. die Größenangaben der Tumore. Umfassend beurteilt werden kann ein Test dann, wenn zu den Ergebnissen eine Vierfeldertafel mit Angabe der richtig und falsch positiven sowie richtig und falsch negativen Befunde vorliegt [27]. Die Qualität der Angaben in diagnostischen Studien zu Krebserkrankungen ist generell als stark verbesserungsbedürftig zu werten, dies wird durch einen aktuellen systematischen Review belegt [28]. Die Publikationsqualität wird bei der methodischen Bewertung der Studien gesondert aufgeführt.

Literatur

22. Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58.209 women with breast cancer and 101.986 women without the disease. *Lancet* 2001;358(9291):1389-99
23. Tyrer J, Duffy SW, Cuzick JA. Breast cancer prediction model incorporating familial and personal risk factors. *Stat Med* 2004;23(7):1111-30
24. Gordis L. *Epidemiologie*. Marburg: Kilian-Verlag; 2001
25. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ* 2006;332(7543):689-92
26. Weiss Ch. *Basisbegriffe medizinischer Statistik*. Berlin: Springer; 2001
27. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC; [STARD Group]. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Fam Pract* 2004;21(1):4-10
28. Mallet S, Deeks JJ, Halligan S, Hopewell S, Cornelius V, Altman DG. Systematic reviews of tests in cancer: review of methods and reporting. *BMJ* 2006; 333:413-9

4. Ergebnisse

4.1 Stellenwert der ergänzenden Mammasonographie im Rahmen des Brustkrebs-Screenings

4.1.1 Ergebnisse der Literaturrecherche

Die Suche in DARE ergab 11 Treffer, die Suche in Cochrane 24. Keines der Abstracts dieser Publikationen erfüllte inhaltlich die Einschlusskriterien.

Rechercheergebnisse PubMed

PubMed – Ergebnisse gesamt 1.1.2000-2.3.2007	n= 731
Nach Screening der Abstracts geprüfte Volltexte	n= 16
Reviews	n= 7
Systematische Reviews	n= 0
Einzelstudien zur Mammasonographie ergänzend zur Mammographie bei asymptomatischen Frauen	n= 9
Davon Studien, die die Einschlusskriterien erfüllten	n= 6

4.1.2 Aufbereitete Evidenz

Die sieben identifizierten Reviews wurden ausgeschlossen, da sie nicht den Kriterien systematischer Reviews entsprachen und/oder nicht die fokussierte Fragestellung beantworteten (siehe 6.1.2).

4.1.3 Evidenz von Einzelstudien

a) Studien- und Publikationsqualität

Für den Zeitraum von 1. Januar 2000 bis 2. März 2007 wurden neun unizentrische Kohortenstudien identifiziert. Von diesen erfüllten sechs prospektive Kohortenstudien die Einschlusskriterien (siehe Evidenztabellen 5.1.1 Evidenz von Einzelstudien und 5.1.2 ausgeschlossene Studien).

Ein konsekutiver Einschluss von Patientinnen war in fünf von sechs Studien den Angaben zu entnehmen. Der Referenzstandard, die histologische Sicherung der Befunde, wurde für als maligne oder nicht bestimmbar klassifizierte Befunde fast durchgängig angewendet.

Als normal oder benigne klassifizierte Befunde wurde allerdings in den vorliegenden Studien nicht konsequent histologisch gesichert oder alternativ entsprechend nachbeobachtet.

Lediglich Kolb et al., 2002 gab eine Nachbeobachtung für alle benignen Befunde über ein Jahre lang (bis zum nächsten Screening) an, jedoch nicht für die Normalbefunde. Kaplan et al., 2001 wiesen auf ein Follow-up für einen Teil der als benige klassifizierten Befunde hin. Den übrigen Studien waren keine Angaben auf ein Follow-up zu entnehmen.

Ein nicht konsekutiver Einschluss von Patientinnen und/oder ein nicht durchgängig angewendeter Referenzstandard führte in allen Studien zu einer Klassifizierung in die Evidenzkategorie 3b.

Hinsichtlich der Publikationsqualität ist festzustellen, dass nur bei fünf der sechs Studien die falsch positiven Befunde so ausgewiesen waren, dass der positive prädiktive Wert berechnet werden konnte. Die Angabe des medianen Alters fehlte in zwei Studien. Der relative Anteil der zusätzlich zur Mammographie entdeckten Karzinome war in fünf der sechs Studien angegeben.

b) Inhaltliche Ergebnisse

Gruppe untersuchter Frauen

Die systematische Suche nach Brustkrebs-Screening-Studien, die zusätzlich zur Mammographie auch die Mammasonographie einsetzen, erbrachte ausschließlich Studien, in denen die Mammasonographie als Zusatzuntersuchung nach vorausgegangener negativer Mammographie angewendet wurde.

Aus der Gruppe von Frauen mit negativer Mammographie wurden weiterhin – mit Ausnahme der Studie von Leconte et al., 2003 – nur Frauen mit dichtem Brustdrüsengewebe mammasonographiert.

Die folgenden Ergebnisse beziehen sich demzufolge aus der Gesamtpopulation untersuchter Frauen im Zeitraum nur auf Frauen mit einer negativen Mammographie und – mit Ausnahme einer Studie – dichtem Brustdrüsengewebe.

Der Anteil dieser Frauen in Bezug auf alle im Zeitraum untersuchten Frauen war bei zwei Studien angegeben und betrug 36% bzw. 48% aller untersuchter Frauen (Kaplan et al., 2001 und Corsetti et al., 2006). Die Größe der untersuchten Studienpopulationen betrug zwischen $n = 1.517$ und $n = 8.103$, im Mittel $n = 4.598$.

Altersverteilung

Das mediane Alter war bei vier Studien angegeben und lag – bei Altersspannen von jeweils mehr als 30 Jahren – zwischen 47,6 Jahren und 60,7 Jahren. Die beiden Studien mit fehlender Angabe des medianen Alters nannten als Altersspanne 25-96 Jahre und 35-87 Jahre (Corsetti et al., 2006; Kaplan et al., 2001).

Karzinomdiagnosen nach Brustdichte-Kategorien

Lediglich in einer Studie waren Frauen aller Brustdichte-Kategorien eingeschlossen (Leconte et al., 2003).

Die weiteren Studien wendeten die Mammasonographie bei Frauen mit dichter Brust an. In zwei Studien wurden mammographische Befunde der Brustdichte BI-RADS-Klassen 3 bis 4 und in drei Studien die BI-RADS-Klassen 2 bis 4 untersucht.

In vier der sechs Studien erfolgte eine Zuweisung der durch Mammasonographie diagnostizierten Karzinome zu der BI-RADS-Kategorie der Brustdichte. Frauen mit Brustdichte BI-RADS 3 und/oder 4 wiesen jeweils den höchsten Anteil an durch Mammasonographie diagnostizierter Mammakarzinome auf.

Leconte et al. diagnostizierten insgesamt 16 Karzinome, davon 11 (72% relativ) bei Frauen der Brustdichte 3 und 4 und fünf bei Frauen der Brustdichte 1 und 2.

Buchberger et al. diagnostizierten 36 Karzinome bei Brustdichte 3 und 4 (0,3% und 1,1% absolut), während zwei bei Brustdichte 2 (0,4%) gefunden wurden. Crystal et al. fanden bei Frauen der Brustdichte 2 kein Karzinom und bei Frauen der Klasse BI-RADS Klasse 3 und 4 0,4% bzw. 0,3%. Kaplan et al. diagnostizierten 0,11% Karzinome in der Klasse BI-RADS 2, und 0,27% bzw. 0,25% in der Klasse BI-RADS 3 und 4.

Drei Studien berichteten von einem retrospektiv gefundenen Anteil von mind. 50% Frauen mit hohem familiärem Risiko bei durch Mammasonographie entdeckten Karzinomen.

Anteil und Größe der durch Mammasonographie zusätzlich diagnostizierten Karzinome

Der relative Prozentsatz zusätzlich durch Mammasonographie entdeckter Karzinome in Bezug auf die Gesamtzahl entdeckter Karzinome war in fünf Studien angegeben und lag im Mittel bei 22% (15%-34%). Der absolute Anteil in Bezug auf die untersuchte Gesamtpopulation war für alle Studien angegeben und lag im Mittel bei 0,36% (0,23%-0,46%). Die Größe der diagnostizierten malignen Tumoren lag bei fünf der sechs Studien bei < 2 cm, lediglich zwei Studien berichteten auch von Karzinomen größer 2 cm, die mediane Größe der entdeckten Karzinome lag bei allen Studien zwischen 9 mm und 1 cm.

Testgüte

Der positive prädiktive Wert für die Mammasonographie in Hinblick auf die Erkennung zusätzlicher maligner Befunde war für vier der sechs Studien angegeben oder ermittelbar und lag im Mittel bei 18,8% (2% bis 33%). Dies bedeutet, dass bei 67% bzw. bei bis zu 98% der als positiv für Malignität klassifizierten Befunde in der Folge kein Karzinom nachgewiesen wurde. Die stark variierenden positiven prädiktiven Werte lassen sich wesentlich auf die Verwendung unterschiedlicher sonographischer Malignitätskriterien und Bewertungskategorien zurückführen (siehe nächster Abschnitt).

Ein Follow-up im Sinne einer Nachbeobachtung zwischen zwei Screening-Untersuchungen war nur in einer Studie für alle als benigne klassifizierten Befunde angegeben (Kolb et al., 2002). Für die als normal Befundeten gaben Kolb et al. jedoch kein Follow-up an. Bei einer weiteren Studie erfolgte ein Follow-up nur für einen geringen Anteil der benignen Befunde (Kaplan et al., 2001). Die jeweils angegebene Sensitivität von 100% ebenso der angegebene negative prädiktive Wert von 100% sind deshalb nicht als valide zu werten.

Die Spezifität konnte für keine Studie ermittelt werden.

Mammasonographische Bewertungskategorien

Der positive prädiktive Wert der mammasonographischen Karzinomentdeckung lässt sich wesentlich auf die verwendeten sonographischen Bewertungskriterien und Bewertungskategorien zurückführen.

Kaplan et al., 2001 verwendeten eine 2-armige Kategorisierung, nämlich die einfache Einteilung in negativ und positiv. Alle positiven Befunde wurden als potentiell suspekt gewertet. Im Unterschied zu den anderen Studien werteten Kaplan et al. auch Zysten als positiv, wenn sie größer als 1 cm waren. Sie erzielten einen positiven prädiktiven Wert von 2%.

Drei Studien (Buchberger et al., 2000; Kolb et al., 2002 und Leconte et al., 2003), verwendeten eine Einteilung der Befunde in drei Kategorien, die Kategorien waren jedoch unterschiedlich definiert. Die Kategorien 1, 2 und 3 lauteten ‚normal‘, ‚benigne‘ und ‚suspekt‘ (Kolb et al.), bzw. ‚gutartig‘, ‚nicht bestimmbar‘ und ‚maligne‘ (Buchberger et al.). Hier wurden positive prädiktive Werte von 10,3% (Kolb et al.) und von 30% (Buchberger et al.) erzielt. Für die Studie von Leconte konnten die Werte nicht ermittelt werden.

Crystal et al., 2003 verwendeten eine vierarmige Kategorisierung und erzielten damit einen positiven prädiktiven Wert für Malignität von 36,8%. Im Unterschied zu Buchberger et al. waren die ‚nicht bestimmbar‘ Befunde in diese Berechnung einbezogen. Nur eine Studie von 2006 (Corsetti et al.) wendete die fünf Kategorien umfassende BI-RADS- Klassifikation für mammasonographische Befunde an. Für diese Studie war der positive prädiktive Wert jedoch nicht ermittelbar.

Biopsie-Raten aufgrund von Mammasonographien

Die aus auffälligen mammasonographischen Befunden resultierende Biopsie-Rate betrug 2,5% (Crystal et al., 2003) bis 7,5% (Corsetti et al., 2006) in Bezug auf das Gesamtkollektiv untersuchter Frauen. Der positive prädiktive Wert für diese Biopsien im Hinblick auf die Entdeckung von Malignomen lag zwischen 3% und 21,5%. Der Unterschied im Hinblick auf den positiven prädiktiven Wert der mammasonographischen Kategorisierung erklärt sich aus der Tatsache, dass Biopsien teilweise auch von Befunden erhoben wurden, die als ‚nicht bestimmbar‘ oder als ‚gutartig‘ eingestuft wurden. Das bedeutet klinisch, dass 88,5% bis 97% der invasiven, diagnosesichernden Eingriffe nicht zu einer Karzinomdiagnose führten (für diese Berechnung wurden minimal-invasive und offene Biopsien nicht getrennt betrachtet).

4.1.4 Diskussion

a) Methodische Diskussion

Für die vorliegende Fragestellung ‚zusätzlicher Nutzen der Mammasonographie im Rahmen des Brustkrebs-Screenings‘ wurde eine sensitive Suchstrategie gewählt (siehe 3.2.1). Die Definition von ‚Screening-Mammographie‘ wurde sehr weit gefasst. Dennoch konnten nur wenige, unizentrische Kohortenstudien identifiziert werden. Diese wurden allerdings alle prospektiv geplant und durchgeführt. Die Studienpopulationen wiesen hohe Altersspannen auf. Die Ergebnisse können deshalb nicht direkt auf eingeladene Frauen eines organisierten Mammographie-Screenings im Alter von 50-69 Jahren übertragen werden.

Aufgrund der Anwendung unterschiedlicher sonomorphologischer Bewertungskriterien und Kategorisierungen sind die Studien nur eingeschränkt in ihren Ergebnissen vergleichbar. Für die umfassende Beurteilung der Testgüte der Intervention Mammasonographie fehlte in den meisten Studien zur Berechnung der Sensitivität, Spezifität und des negativen prädiktiven Werts ein entsprechendes Follow-up. Darüber hinaus war der wesentliche Endpunkt in Bezug auf die klinische Relevanz, die Mortalitätsrate an Brustkrebs bei keiner Studie angegeben. Aufgrund der Studienlage ist deshalb nur eine eingeschränkte Bewertung der Intervention ‚zusätzliche Mammasonographie‘ beim Brustkrebs-Screening möglich.

b) Inhaltliche Diskussion

Die Wirksamkeit der Mammasonographie als diagnostisches Verfahren wurde überwiegend bei Frauen mit Brustdichte der BI-RADS-Klasse 2 bis 4 und grundsätzlich nur bei vorausgegangener negativer Screening-Mammographie untersucht. Aufgrund der hohen Sensitivität der Mammographie bei Brustdichte der BI-RADS-Klasse 1 und der im Vergleich schlechteren Sensitivität bei BI-RADS-Klasse 2 bis 4 ist die Wahl der Studienpopulationen begründbar. Frauen mit dichter Brust haben im Vergleich zu Frauen mit nicht dichter Brust insgesamt ein erhöhtes Risiko, an Brustkrebs zu erkranken [29].

Die Ergebnisse zeigen, dass durch die zusätzliche Intervention Mammasonographie (durchgeführt mit einer Schallkopffrequenz ab 5 Mhz) nach negativer Screening-Mammographie vorwiegend invasive Karzinome in

der Gruppe dieser Frauen im Mittel bei 0,36% absolut diagnostiziert werden können. Zum Vergleich: im populationsbezogenen Screening werden durch Mammographie allein im Mittel bei 0,4% -0,9% aller untersuchten Frauen Karzinome diagnostiziert [30].

Der überwiegende Anteil der Karzinome wurde in den BI-RADS-Gruppen 3 und 4 diagnostiziert.

Die durch Mammasonographie zusätzlich diagnostizierten Karzinome sind diesen Studien zufolge im Mittel nicht größer als die durch Screening-Mammographie diagnostizierten Karzinome.

Berg (2004) gibt in einem Review zum Einsatz der Mammasonographie bei Frauen mit dichtem Brustgewebe ähnliche Ergebnisse an. Der Review war ausgeschlossen worden, weil keine systematische Suchstrategie angegeben war, Berg analysierte jedoch fünf der sechs identifizierten Studien (Buchberger et al.; Kaplan et al.; Kolb et al.; Crystal et al. und Leconte et al.) sowie eine Studie von Gordon et al., 1995. Sie berechnete für die Studienpopulationen gepoolte Werte zur zusätzlichen Karzinomentdeckung mit 0,35% und gab an, dass 94% der durch Mammasonographie entdeckten Karzinome invasiv waren, mit einem mittleren Durchmesser von 9-11 mm. Sie ermittelte, dass über 90% der Frauen mit mammasonographisch entdeckten Karzinomen Brustdichte der BI-RADS-Kategorie 3 und 4 aufwiesen.

Der Beitrag der Mammasonographie in Bezug auf die Karzinomentdeckung des Brustkrebs-Screenings ist demzufolge überwiegend für Frauen der BI-RADS-Kategorie 3 und 4 als relevant zu werten. Ob die verbesserte Karzinomentdeckung auch einen relevanten klinischen Nutzen in Bezug auf den Endpunkt Brustkrebsmortalität erbringt, kann aufgrund der vorliegenden Studien nicht beantwortet werden.

Für die Bewertung des Schadens durch die zusätzliche Mammasonographie bei vorausgegangener negativer Mammographie und dichtem Brustdrüsengewebe gilt Folgendes: Die Biopsieraten für die vorliegenden Studien sind mit 2,5%-7,5% im Vergleich zu den Biopsieraten aufgrund von Screening-Mammographien deutlich erhöht. Hier werden Biopsieraten von ca. 1-2% berichtet [31;32; 33]. Der positive prädiktive Wert der Biopsien (im Mittel 12,6%) ist darüber hinaus deutlich geringer als der von Screening-Mammographien (im Mittel ca. 38%, [32-34]), d.h. fast drei Mal so viel Frauen müssen sich einer Biopsie unterziehen, bis ein Karzinom diagnostiziert wird. Auch wenn sonographische sichtbare Befunde einfach minimal-invasiv biopsiert werden können, ist bei der Beurteilung des Schadens die psychologische Belastung der Frauen vor und während einer Biopsie aufgrund eines falsch positiven mammasonographischen Ergebnisses zu bedenken.

Auffallend ist die Heterogenität der mammasonographischen Bewertungskriterien, die entsprechend die Biopsieraten beeinflusst. Dadurch sind die Studien nur eingeschränkt vergleichbar. Der Einfluss unterschiedlich erfahrener Befunder auf das Ergebnis kann auch aufgrund dieser Tatsache kaum quantifiziert werden.

Aufgrund der nicht mit Sicherheit angebbaren Raten falsch-negativer Befunde kann eine sichere Aussage darüber, ob durch eine zusätzliche Mammasonographie alle durch Mammographie nicht entdeckten Karzinome in der Gruppe der Frauen mit dichter Brust erkannt werden können, nicht getroffen werden.

Lediglich eine prospektiv-randomisierte Studie mit dem Endpunkt Mortalität könnte zeigen, ob die zusätzlich entdeckten Karzinome einen Überlebensvorteil bringen oder eher zu einer ‚Überdiagnostik‘ führen. Eine randomisierte Studie mit diesem Endpunkt existiert allerdings nicht.

Die einzige derzeit laufende randomisierte Studie untersucht die Mammasonographie im Hinblick auf ihre Testgüte als Screening-Untersuchung für Hochrisikopatientinnen mit dichter Brust und wurde 2003 von Berg [34] initiiert. Die Ergebnisse sind zurzeit noch nicht verfügbar.

Im Hinblick auf eine Übertragbarkeit auf ein populationsbezogenes Mammographie-Screening, wie es zurzeit z.B. in Deutschland durchgeführt wird, kann abschließend Folgendes festgestellt werden:

1. Die vorliegenden Studien geben einen Hinweis auf einen zusätzlichen Nutzen der Mammasonographie nach negativer Mammographie im Rahmen eines Brustkrebs-Screenings bei Frauen mit dichter Brust (v.a. BI-RADS Klassen 3 und 4) für die Diagnostik vorwiegend invasiver Karzinome, die im Median bei 9-11 mm liegen. Da die untersuchten Altersgruppen eine sehr große Altersspanne aufwiesen und durchgehend auch jüngere Frauen mit einbezogen waren, wird der Effekt für ein organisiertes Screening-Programm, das ausschließlich für die eingeladenen Frauen zwischen 50 und 69 Jahren gilt, möglicherweise geringer ausfallen. Deshalb sind prospektive Validierungsstudien im Rahmen eines realen Screening-Programms erforderlich.

2. Validierungsstudien sollten für eine multizentrisch rekrutierte Population neben dem positiven prädiktiven Wert auch die Sensitivität, Spezifität und den negativen prädiktiven Wert der Intervention Mammasonographie angeben.
3. Im Hinblick auf den Schaden durch die Intervention Mammasonographie führt eine hohe Rate an falsch-positiven sonographischen Befunden zu mehr als doppelt so hohen Raten an Biopsien im Vergleich zu Biopsien aufgrund von Screening-Mammographien. Validierungsstudien sollten ein einheitliches Bewertungssystem aufweisen, vorzugsweise das BI-RADS-System, um den Einsatz von Biopsien genau begründbar zu machen.
4. Wie für die Bewertung des Mammographie-Screenings sind auch für die Mammasonographie Langzeitstudien mit randomisierter Zuteilung der Patientinnen erforderlich, um den Effekt einer zusätzlich durchgeführten Mammasonographie nach negativer Mammographie im Hinblick auf die Brustkrebsmortalität zu belegen.

Literatur

29. Tyrer J, Duffy SW, Cuzick JA. Breast cancer prediction model incorporating familial and personal risk factors. *Stat Med* 2004 Apr 15;23(7):1111-30
30. International Agency for Research on Cancer (IARC). Screening for breast cancer. IARC Pr.; 2002
31. Kwek BH, Lau TN, Ng FC, Gao F. Non-consensual double reading in the Singapore Breast Screening Project: benefits and limitations. *Ann Acad Med Singapore* 2003;32(4):438-41
32. Morton MJ, Whaley DH, Brandt KR, Amrami KK. Screening mammograms: interpretation with computer-aided detection--prospective evaluation. *Radiology* 2006;239(2):375-83
33. Cupples TE, Cunningham JE, Reynolds JC. Impact of computer-aided detection in a regional screening mammography program. *AJR* 2005;185(4):944-50
34. Berg WA. Rationale for a trial of screening breast ultrasound: American College of Radiology Imaging Network (ACRIN) 6666. *AJR* 2003;180(5):1225-8

4.2 Stellenwert der axillären Sonographie bei klinisch unauffälliger Axilla

4.2.1 Ergebnisse der Literaturrecherche

Zu diesem Thema ergab die Suche in DARE und Cochrane keine Treffer.

Rechercheergebnisse PubMed

PubMed-Ergebnisse gesamt 1.1.2000-2.3.2007	n= 226
Nach Screening der Abstracts geprüfte Volltexte	n= 38
Reviews	n= 3
Systematische Reviews	n= 1
Einzelstudien zur präoperativen Evaluierung der Axilla mittels Sonographie bei Patientinnen ohne neoadjuvante Chemotherapie	n= 34
Davon Studien, welche die Einschlusskriterien erfüllen	n= 8

4.2.2 Aufbereitete Evidenz

Es wurde ein systematischer Review identifiziert: Alvarez S, Anorbe E, Acorta P et al., 2006; 'Role of Sonography in the Diagnosis of Axillary Lymph Node Metastasis in Breast Cancer: A Systematic Review' (siehe 6.2.1). Dieser Review schloss nach systematischer Literaturrecherche in Medline von Januar 1980 bis März 2004 und Auswahl der Studien nach definierten Kriterien 16 Studien ein, von diesen waren 10 Studien vor 2000 durchgeführt worden. Es wurden 14 prospektive und zwei retrospektive Studien eingeschlossen. Acht

der Studien hatten eine konsekutive Patientinnenrekrutierung angegeben. Goldstandard war die definitive histologische Aufarbeitung der Lymphknoten nach der axillären Lymphonodektomie oder der operativen Entnahme der Sentinel-Lymphknoten. Der Review wurde aufgrund der Tatsache, dass acht Studien eingeschlossen wurden, in denen keine konsekutive Patientinnenrekrutierung erfolgt war, nach der Oxford-Klassifikation formal mit dem Evidenzgrad 3a bewertet.

Als sonographisches Kriterium für einen malignen axillären Lymphknoten war neben Größe und Morphologie auch das Ergebnis minimal-invasiver Feinnadel- oder Core-Biopsie unter sonographischer Kontrolle zugelassen.

Folgende Ergebnisse wurden berichtet: bei nicht-palpablen axillären Lymphknoten und Größe als Malignitätskriterium lag die Sensitivität für das Erkennen karzinomatös befallener Lymphknoten zwischen 48,8% und 87,1% und die Spezifität zwischen 55,6% und 87,3%. Bei Morphologie als Malignitätskriterium lag die Sensitivität zwischen 26,4% und 75,9% und die Spezifität zwischen 88,4% und 98,1%. Die Autoren kommen zu dem Schluss, dass eine negative Axillasonographie bei Brustkrebs karzinomatös befallene Lymphknoten nicht mit hinreichender Sicherheit ausschließt.

Ausgeschlossen wurden zwei narrative Reviews (siehe 6.2.4).

4.2.3 Evidenz von Einzelstudien

Aufgrund der systematischen Literaturrecherche wurden insgesamt 34 Einzelstudien zum Thema identifiziert, von denen nach Studium der Volltexte acht Studien eingeschlossen wurden.

Die ausgeschlossenen Einzelstudien enthielten entweder keine Angaben zur Intervention Axillasonographie alleine oder die Ergebnisse waren nicht auf klinisch unauffällige axilläre Befunde bezogen (siehe 6.2.2 Evidenz von Einzelstudien und 6.2.3 ausgeschlossene Studien). Eine der eingeschlossenen Studien (Deurloo et al., 2003) ist auch in dem Review von Alvarez et al., 2006 enthalten.

a) Studien- und Publikationsqualität

Es wurden acht Kohortenstudien mit ausgewiesenen Angaben für Patientinnen mit klinisch negativem axillären Befund identifiziert.

Fünf dieser Studien wurden prospektiv durchgeführt (Rajesh et al., 2002; Deurloo et al., 2003; Esen et al., 2004; Mathijssen et al., 2006; Nori et al., 2007) mit einer mittleren Studienpopulation von n= 152 (n= 82 bis n= 265).

In drei Studien wurden retrospektive Analysen erstellt (van Rijk et al. 2006; Bedrosian et al. 2003; Podkrajsek et al. 2005 (Studienpopulation im Mittel n= 366 (n= 165-n= 726)).

Da alle Studienpatientinnen im Rahmen ihrer Primärtherapie eine definitive axilläre Operation erhielten, konnten für alle Befunde der positive und der negative prädiktive Wert sowie die Sensitivität und Spezifität angegeben werden, d.h. der Referenzstandard wurde durchgehend vollständig angewendet. Obwohl die sonographische Beurteilung axillärer Lymphknoten auch in Studien vor 2000 dokumentiert ist (siehe z.B. systematischer Review von Alvarez et al., 2006), konnten die Studien aufgrund differierender Malignitätskriterien nicht als Validierungsstudien bewertet werden. Wurden die Patientinnen für die Studien konsekutiv rekrutiert, wurde deshalb ein Evidenzgrad von 2b vergeben. Die Angabe war bei drei Studien zu entnehmen (Rajesh et al.; Deurloo et al.; van Rijk et al.). Fehlte die Angabe oder war aus der Studienbeschreibung ersichtlich, dass es sich um nicht-konsequente Untersuchungen handelte, wurde der Evidenzgrad 3b vergeben. Die Verblindung zum Palpationsbefund wurde bei grundsätzlich vorausgehendem negativem Befund als Klassifikationsmerkmal ignoriert.

Hinsichtlich der Publikationsqualität der Studien ist anzumerken, dass die Angaben zum Anteil dargestellter Lymphknoten bei 3 von 8 Studien fehlten, ebenso waren die Angaben zu Anzahl und Erfahrung der Befunder in nur 4 Studien gegeben.

b) Inhaltliche Ergebnisse

Bei negativem klinischen Axillabefund wurden – bis auf die Studien von Rajesh et al. – überwiegend T1-Tumoren beurteilt (67-80%).

Testgüte

Der negative prädiktive Wert für die Erkennung karzinompositiver Lymphknoten bei klinisch negativer Axilla war im Hinblick auf prospektiv und retrospektiv erhobene Daten nicht unterschiedlich und lag im Mittel bei 0,76 (0,64-0,93). Die Sensitivität und Spezifität lagen im Mittel bei 47,6% (26,4-94,45%) und 88,3% (82%-98,7%). Der positive prädiktive Wert wurde zwischen 0,53 und 0,95 angegeben.

Sonographische Bewertungskriterien

Die für die Beurteilung der axillären Lymphknoten verwendeten sonographischen Malignitätskriterien waren unterschiedlich. Nur das Kriterium Hypoechogenität wurde in allen Studien genannt.

Das Kriterium asymmetrischer Kortex wurde in fünf Studien einbezogen. Für die Definition eines asymmetrischen Kortex gaben z.B. Esen et al. an, dass die breiteste Stelle > 50% der dünnsten Stelle betragen sollte, während Bedrosian et al. und Rajesh et al. nur ‚exzentrische Kortexverdickung‘ angaben. Van Rijk et al. verwendeten für die zweite Hälfte ihrer in die Studie einbezogenen sonographischen Untersuchungen nur noch das Merkmal ‚Kortex größer 2,3 mm‘.

Die Malignitätskriterien differierten auch im Hinblick auf die Zunahme des maximalen Lymphknotendurchmessers, dieses Kriterium wurde in 3 Studien genannt. Der cut-off war 5 mm (Deurloo et al.; van Rijk et al.) und 3 mm (Mathijssen et al.). In zwei Studien (Esen et al. und Podkrajsek et al.) wurde der periphere Dopplerflow als weiteres Merkmal genannt (siehe 6.2.2).

Testgüte in Abhängigkeit weiterer Parameter

Esen et al. erreichten bei in 99% dargestellten Lymphknoten und sehr detailliert angegebenen Malignitätskriterien einen negativen prädiktiven Wert von 0,93 und eine Sensitivität von 94,45% unter Einbeziehung des peripheren Dopplerflows.

Nori et al. gaben als Malignitätskriterium unter anderen ebenfalls die Analyse des peripheren Dopplerflows an. Sie konnten bei jeder Untersuchung Lymphknoten darstellen, erreichten aber einen schlechteren negativen prädiktiven Wert von 0,77 bei einer Sensitivität von 45,2%. 20 von 23 falsch-negativ beurteilten Axillae in dieser Studie wiesen allerdings nur Mikrometastasen in den operativ entfernten Lymphknoten auf. Die Lymphknoten der 3 weiteren Axillae waren makrometastatisch mit zwei, vier und acht Lymphknoten befallen. Die beiden genannten Studien enthielten keine Angabe darüber, ob ein konsekutiver Einschluss der Patientinnen erfolgt war. Rajesh et al. untersuchten 45 konsekutive Patientinnen mit klinisch unauffälligem axillären Befund. Sie konnten ebenfalls bei jeder Patientin Lymphknoten darstellen und erreichten einen negativen prädiktiven Wert von 0,87 bei einer geringen Sensitivität von 38%.

Deurloo et al., die alle Patientinnen im Zeitraum konsekutiv untersuchten, erreichten bei in 35% der Untersuchungen dargestellten Lymphknoten einen negativen prädiktiven Wert von 0,64 und eine Sensitivität von 40,5%. Hinsichtlich der Vorerfahrung gaben Deurloo et al. an, dass die sieben Befunder jeweils lediglich mindestens fünf axilläre Sonographien vor der Studie durchgeführt hatten. Bei Rajesh et al., Esen et al. und Nori et al. wurde die Untersuchung jeweils nur von einem Befunder vorgenommen.

Das größte Kollektiv wurde retrospektiv von van Rijk et al. als Folgestudie zu der Erhebung von Deurloo et al. untersucht. Eine konsekutive Rekrutierung war angegeben, die Anzahl und Erfahrung der Untersucher war nicht angegeben. Hier wurde ein negativer prädiktiver Wert von 0,68 bei einer Sensitivität von 34,7% erreicht.

Minimal-invasive Diagnoseverfahren während der Mammasonographie

In fünf Studien wurde an sonographisch suspekten Lymphknoten während der Untersuchung eine Feinnadelpunktion vorgenommen, Nori et al. führten bei suspekten Lymphknoten eine Stanzbiopsie durch. Die zytologische bzw. im letzteren Fall die histologische Sicherung des sonographischen Befundes zeigte durchgängig eine Spezifität von 91-100% bei einer Sensitivität von 61,7-91,6%. Lediglich in der retrospektiven Studie von Bedrosian et al. wurden nur 25% erreicht. Der negative prädiktive Wert war der Axillasonographie vergleichbar.

4.2.4 Diskussion

a) Methodische Diskussion

Für die Beantwortung der Fragestellung zur diagnostischen Sicherheit der axillären Sonographie wurden aufgrund der systematischen Literaturrecherche ein aktueller systematischer Review (Alvarez et al., 2006) und acht Einzelstudien berücksichtigt.

Obwohl in dem systematischen Review auch Studien eingeschlossen waren, bei denen keine Angaben zur Sonographie alleine, sondern kombiniert mit einer zytologischen Sicherung vorlagen, konnte der Review als

Bewertungsgrundlage und Vergleich zu den identifizierten Einzelstudien herangezogen werden. Die zusätzliche Intervention Feinnadelpunktion beeinflusst die Sensitivität und den negativen prädiktiven Wert der Axillasonographie alleine kaum, wenn sie nur bei auffälligen Lymphknoten durchgeführt wird. Dies war bei den im Review eingeschlossenen 16 Studien mit Ausnahme einer Studie der Fall.

In Bezug auf das Rechercheergebnis für Einzelstudien (34 Einzelstudien zum Thema) wurden die meisten Studien aufgrund zweier Gründe ausgeschlossen: entweder sie wiesen die Ergebnisse der axillären Sonographie nicht gesondert aus und es war damit keine spezifische Berechnung der weiteren Testgüteparameter möglich oder es wurden auch Patientinnen mit klinisch auffälliger Axilla in Kollektiven mit erhöhter Prävalenz befallener Lymphknoten untersucht.

Aufgrund der Tatsache, dass die operative Entfernung von Lymphknoten an der Axilla und ihre histologische Aufarbeitung ein Standardeingriff im Rahmen der Primäroperation von Brustkrebs ist, war der Referenzstandard durchgängig gewährleistet.

Ein Nachteil für die Beurteilung sowohl bei den im Review eingeschlossenen Studien als auch bei den zusätzlich identifizierten Einzelstudien ist, dass nicht für alle eine konsekutive Patientinnenrekrutierung angegeben war. Des Weiteren sind die erzielten Ergebnisse aufgrund unterschiedlicher sonographischer Malignitätskriterien nur eingeschränkt vergleichbar.

Die diagnostische Sicherheit der Axillasonographie kann deshalb nicht abschließend bewertet werden.

b) Inhaltliche Diskussion

In den vorliegenden Studien wurden – bis auf die Studie von Rajesh et al. – überwiegend axilläre Sonographien bei T1-Tumoren durchgeführt. Dadurch wurde die Situation der Brustkrebsfrüherkennung mit geringer Prävalenz an befallenen Lymphknoten gut abgebildet.

Der Anteil dargestellter Lymphknoten korreliert sicher mit der Erfahrung der Befunder. Ein hoher Anteil von Untersuchungen mit sonographisch dargestellten Lymphknoten führte jedoch den vorliegenden Studien nicht immer zu einer hohen Sensitivität der Axillasonographie.

Die Sensitivität der axillären Sonographie weist eine hohe Variabilität auf (26,4-94,5%), ebenso wie der negative prädiktive Wert (0,64-0,93), dies zeigte die aufbereitete Evidenz ebenso wie die zusätzlichen Einzelstudien. Die Vergleichbarkeit der Studien wird erschwert durch die Anwendung unterschiedlicher sonographischer Malignitätskriterien. In der Gesamtbeurteilung der Studienlage kann die unauffällige Axillasonographie bei negativem klinischem Befund karzinomatös befallene Lymphknoten derzeit nicht mit so großer Sicherheit ausschließen, dass auf eine weitere operative Abklärung verzichtet werden kann. Für eine Validierung oder Revision dieses Ereignisses sind konsekutive Studien an mehreren Zentren erforderlich, bei denen gleiche Malignitätskriterien angewendet werden und eine vergleichbare Expertise der Befunder gegeben ist.

4.3 Nutzen der Mammasonographie zur Karzinomdiagnostik bei auffälligem Brustbefund

4.3.1 Ergebnisse der Literaturrecherche

Für diese Fragestellung wurde in DARE und Cochrane ein Abstract von 1999 identifiziert, das die inhaltlichen Einschlusskriterien nicht erfüllte.

Rechercheergebnisse PubMed

PubMed 1/1999- 10.4.2007	n= 361
Nach Screenen der Abstracts geprüfte Volltexte	n= 27
Reviews	n= 3
Systematische Reviews	n= 0
Einzelstudien ab 2000	n= 18
zusätzliche Einzelstudien 1999	n= 6
Eingeschlossene Studien	n= 11

4.3.2 Aufbereitete Evidenz

Es wurde kein systematischer Review identifiziert. Ausgeschlossen wurden drei narrative Reviews (siehe 6.1.3).

4.3.3 Evidenz von Einzelstudien

a) Studien- und Publikationsqualität

Die systematische Literaturrecherche zum Stellenwert der Mammasonographie bei auffälligem Brustbefund erbrachte auch bei Ausweitung des Recherchezeitraums um ein Jahr insgesamt wenige Studien, die die Mammasonographie als diagnostisches Verfahren – zumindest in Bezug auf einzelne Parameter - getrennt von der Mammographie bewerteten.

Es wurden sieben prospektive Kohortenstudien (mittl. Größe der Studienpopulation n= 348 (111-559) und vier retrospektive Kohortenstudien (mittl. n= 168, 103-293) identifiziert (siehe 6.3.1. Evidenz von Einzelstudien und 6.3.2 ausgeschlossene Studien).

In sechs dieser elf Studien wurden Patientinnen konsekutiv rekrutiert. Nur in einer Studie erfolgte die mammasonographische Untersuchung verblindet zur Mammographie (Skaane et al., 1999). Bei jüngerem Alter der Patientin wurde in den Studien die Mammasonographie zum Teil vor der Mammographie vorgenommen. Nicht bei jeder Studie wurden alle Patientinnen sowohl mammographiert als auch mammasonographiert.

Die Anwendung des Referenzstandards, die histologische oder zumindest zytologische Beurteilung der Befunde bzw. eine Nachbeobachtung über mindestens 12 Monate war bei fast allen Studien gegeben. Nur die Studie von Shetty et al. wies kein ausreichendes Follow-up auf. Insgesamt konnte nach Anwendung der Oxford-Kriterien jedoch keine der Studien mit einem Evidenzlevel von 2 bewertet werden. Es wurde jeweils ein Evidenzlevel 3b vergeben.

Die Studien waren in Bezug auf die Studienpopulationen heterogen. In zwei der Studien war jeweils nur ein Anteil von 73% klinisch erhobener, tastbarer Brustbefunde angegeben (Berg et al., 2004 und Rahbar et al., 1999). Cid et al., 2004, bewerteten den diagnostischen Nutzen der Mammasonographie bei durch die Patientin berichtetem Tastbefund an der Brust, der bei der ärztlichen Untersuchung nicht reproduziert werden konnte. Bei den weiteren Studien war für alle Patientinnen von einem klinisch erhobenen Brustbefund auszugehen. Kaiser et al., 2002, betonten, dass es sich bei den durch standardisierte klinische Untersuchung erhobenen Befunden um ‚breast thickening‘, nicht um einen typischen Herdbefund handelte. In den anderen Studien wurde die Art des palpablen Befunds nicht spezifiziert.

In vielen Studien fehlten relevante Angaben. Bei drei Studien war das mediane Alter der Studienpopulation nicht angegeben. In acht der Studien fehlten die Größenangabe bzw. die Angabe des T-Stadiums der durch die Mammasonographie diagnostizierten Tumoren. In drei Studien waren die sonographischen Malignitätskriterien nicht angegeben. Die Anzahl der falsch-positiven Befunde war in zwei Studien nicht angegeben.

Die Sensitivität und Spezifität war insgesamt bei vier der Studien ermittelbar.

Eine altersstratifizierte Auswertung erfolgte bei zwei Studien (Flobbe et al., 2003; Zonderland et al., 1999). Diese Auswertung bezog sich auf den Gewinn an diagnostischer Sicherheit durch den zusätzlichen Einsatz der Mammasonographie. Bei Flobbe et al. war eine klinische Untersuchung und Mammographie vorausgegangen, bei Zonderland et al. nur die Mammographie.

b) Inhaltliche Ergebnisse

Altersverteilung

In sechs von acht Studien mit der entsprechenden Angabe lag das mediane Alter unter 50 Jahren. Insgesamt war es zwischen 39,4 Jahren und 64 Jahren angegeben. Die Altersspannen betragen dabei jeweils mehr als 45 Jahre.

Anteil und Größe an durch Mammasonographie allein diagnostizierten Karzinomen

Die Rate der durch Mammasonographie diagnostizierten Karzinome, die mammographisch okkult waren bzw. die mammographisch als nicht bestimmbar oder als benigne eingestuft wurden, war für sieben Studien

ermittelbar. Aufgrund der Mammasonographie wurden zwischen 6% (n= 8) und 30,9% (n= 17) Karzinome zusätzlich diagnostiziert.

Berg et al. gaben die Größe dieser Karzinome zwischen 4 und 90 mm (im Median 27 mm) an, Georgian-Smith et al. gaben eine mediane Größe von 1,7 cm an. Skaane et al. 1999 stratifizierten nach Tumoren kleiner und größer 2 cm und fanden nur für T2-Tumore (größer 2 cm) einen signifikanten zusätzlichen Nutzen der Mammasonographie.

Cid et al. gaben bei unauffälliger ärztlicher Palpation aber von der Patientin berichtetem auffälligem Befund ein durch die Mammasonographie diagnostiziertes Karzinom von insgesamt drei im Gesamtkollektiv entdeckten Karzinomen an. Aufgrund der Mammographie wurden in diesem Kollektiv zwei Karzinome entdeckt. Alle drei Karzinome waren kleiner als ein Zentimeter.

Testgüte

Die Sensitivität der Mammasonographie war für vier Studien ermittelbar und betrug zwischen 83% und 100%. Die Spezifität war bei diesen Studien ebenfalls angegeben und betrug zwischen 34% und 95,5%. Der positive prädiktive Wert war nur bei einer Studie angegeben und betrug 60%.

Der negative prädiktive Wert nur für die Mammasonographie war für zwei Studien angegeben und betrug 98% sowie 100%.

Weitere Ergebnisse

Ein Vergleich der Sensitivität der Mammasonographie mit der Mammographie war in fünf Studien gegeben (Berg et al., 2004; Flobbe et al., 2003; Rahbar et al., 1999; Moss et al., 1999; Zonderland et al., 1999) und zeigte eine zwischen 4,7% und 16% höhere Sensitivität für die Mammasonographie. Die Spezifität war vergleichend für drei Studien zu erheben (Berg et al.; Moss et al. und Flobbe et al.) und war für zweimal für die Mammographie um 37% und 4,7% besser, bei Flobbe et al. lag jedoch die Spezifität der Mammasonographie um 3,6% höher als die der Mammographie.

Berg et al., 2004 verglichen prospektiv (n= 111) die diagnostischen Verfahren klinische Untersuchung, Mammographie, Mammasonographie und Magnetresonanztomographie (MRT). Die höchste Sensitivität zeigte die Magnetresonanztomographie, die höchste Spezifität die klinische Untersuchung. Die beste diagnostische Kombination war die klinische Untersuchung kombiniert mit Mammographie und Magnetresonanztomographie.

Shetty et al., 2001 und Rahbar et al., 1999 gaben den aus Mammographie und Mammasonographie kombinierten negativen prädiktiven Wert mit 100% und 96% an.

Flobbe et al., 2003 konnten durch die Mammasonographie in 28% der Patientinnen falsch-positive Befunde aus Mammographie und Palpation zurückstufen.

Altersstratifizierte Auswertungen

Zonderland et al. stratifizierten ihre Auswertungen nach Altersklassen zwischen 30 Jahren und über 70 Jahren, in 10-Jahres-Schritten. Sie fanden für Frauen von 30-50 Jahren eine um 20% verbesserte Sensitivität bei zusätzlichem Einsatz der Mammasonographie und bei Frauen von 50-70 Jahren eine Verbesserung um 6%. Flobbe et al. zeigten ebenfalls eine deutlichere Verbesserung der Sensitivität bei Frauen unter 50 Jahren im Gegensatz zu Frauen über 50 Jahren.

Mammasonographische Bewertungskriterien

Auch bei diesen Studien waren die Bewertungskriterien zum Teil heterogen. In vier Studien (Berg et al., Kaiser et al., Moss et al. und Zonderland et al.) wurde eine vierarmige Kategorisierung benutzt. Kaiser et al. definierten dabei jeden soliden Herdbefund als suspekt. In weiteren vier Studien (Shetty et al., Rahbar et al. und in zwei Studien von Skaane et al.) wurde eine dreiarmlige Bewertung verwendet. Lediglich Flobbe et al. benutzten eine fünfarmige Kategorisierung. In zwei Studien ist kein Bewertungssystem angegeben.

4.3.4 Diskussion

a) Methodische Diskussion

Für die vorliegende Fragestellung des Nutzens der Mammasonographie bei palpablem Befund in Hinblick auf die Karzinomerkenung wurden sehr unterschiedliche Studien identifiziert. Aufgrund der heterogenen Altersverteilung der Studienkollektive sind die Studien nur schwer vergleichbar, da die Prävalenz von Mammakarzinomen innerhalb der Studienkollektive als unterschiedlich angenommen werden muss. Die Vergleichbarkeit wird durch weitere fehlende Angaben und unterschiedliche bzw. fehlende

mammasonographische Bewertungskriterien erschwert. Insgesamt sind die Studien deshalb als zu heterogen zu bewerten, um den Nutzen der Mammasonographie quantitativ valide zu bewerten.

Bei den erzielten Testgüteparametern ist weiterhin zu beachten, dass die Mammasonographie nicht verblindet zur Mammographie erfolgte.

b) Inhaltliche Diskussion

Die vorliegenden Studien zeigen bei auffälligen Brustbefunden in Bezug auf die Karzinomentdeckung eine höhere Sensitivität von 6%-30,9% bei zusätzlichem Einsatz der Mammasonographie verglichen mit der alleinigen Mammographie.

Die Sensitivität der Mammasonographie in Kenntnis des mammographischen Befunds lag zwischen 83% und 100%.

Für die Spezifität der Mammasonographie im Vergleich zur Mammographie zeigten sich sowohl eine schlechtere Spezifität (bis zu -34%) als auch eine bessere Spezifität.

Für die Verbesserung der Früherkennung von Mammakarzinomen ist nur die Diagnostik von Tumoren bis zu 2 cm (T1) von potentielltem Nutzen. Unter den durch Mammasonographie allein diagnostizierten Tumoren bei auffälligem Befund fanden sich auch solche unter zwei Zentimetern. Die vorliegenden Daten zeigen jedoch nicht, dass es sich bei mammographisch nicht erkannten palpablen Karzinom, die aufgrund einer Mammasonographie diagnostiziert werden, vorwiegend um T1-Karzinome handelt. Zusammenfassend hat die Mammasonographie bei auffälligen Brustbefunden eine zur Mammographie ergänzende Rolle, diese ist jedoch nicht spezifisch für die Brustkrebsfrüherkennung.

4.4 Stellenwert der Mammasonographie bei Frauen unter 50 Jahren mit auffälligem Brustbefund

4.4.1 Ergebnisse der Literaturrecherche

Für diese Fragestellung wurden die gleichen Suchbegriffe verwendet wie für Frage 4.3.

Die Suche in DARE und Cochrane ergab keine inhaltlich relevanten Treffer.

Rechercheergebnisse PubMed

PubMed (1/1998-15.4.2007)	n= 401
Nach Screenen der Abstracts geprüfte Volltexte	n= 11
Reviews	n= 0
Einzelstudien	n= 7
Handsuche in Literaturverzeichnissen	n= 4
Ausgeschlossene Studien	n= 0

4.4.2 Aufbereitete Evidenz

Zu dieser Fragestellung wurde kein Review identifiziert.

4.4.3 Evidenz von Einzelstudien

a) Studien- und Publikationsqualität

Für den Stellenwert der Mammasonographie als erstes diagnostisches Verfahren bei symptomatischen Frauen unter 50 Jahren konnten auch nach Ausweitung des Recherchezeitraums bis 1998 nur wenige Studien identifiziert werden. Insgesamt wurden zehn retrospektive Kohortenstudien eingeschlossen (siehe 5.4.1 Evidenz von Einzelstudien). Keine Studie wurde ausgeschlossen.

Keine der Studien gab eine konsekutive Patientinnenrekrutierung an.

Lediglich Houssami et al., 2003, führten eine verblindete Beurteilung von Mammasonographie und Mammographie durch, dies jedoch erst retrospektiv, d.h. die ursprüngliche Bewertung erfolgte nicht verblindet.

Die Studien wurden deshalb bis auf eine Ausnahme mit dem Evidenzgrad 3b bewertet.

Für eine Studie wurde ein Evidenzlevel von 4 vergeben, weil kein Referenzstandard (weitgehend keine Nachbeobachtung) angewendet wurde, auch wenn dies klinisch bei Adoleszenten – in der Studie waren Frauen von 9-19 Jahren eingeschlossen – erklärbar ist.

Bezüglich der Qualität der Angaben ist festzustellen, dass in der Studie von Morrow et al. nur der Anteil falsch positiver Befunde berechnet werden konnte.

b) Inhaltliche Ergebnisse

Altersverteilung

Die obere Altersgrenze der in den vorliegenden Studien eingeschlossenen Patientinnen variierte zwischen weniger als 20 Jahre und bis 55 Jahre. Das mediane Alter war in sechs Studien angegeben und lag zwischen 24,8 und 44,9 Jahren. In insgesamt sieben der 10 Studien waren ausschließlich Frauen unter 40 Jahren eingeschlossen.

In einer Studie wurden Frauen aller Altersklassen untersucht, aber stratifiziert nach Alter ausgewertet, hier wurden die Altersgruppen bis 49 Jahre berücksichtigt.

Testgüte

Die Sensitivität der Mammasonographie war bei sieben Studien angegeben und betrug im Mittel 81,6% (67%-95%). Die Spezifität war nur bei zwei Studien (Houssami et al.) angegeben und betrug 88% und 96,6%. Der positive und negative prädiktive Wert war nur in einer Studie ausgewiesen bzw. berechenbar (Malik et al., 2006) und betrug 75% bzw. 100%. Malik et al. zeigten zusätzlich eine höhere Sensitivität der Mammasonographie für die Erkennung benigner wie für die Erkennung maligner Befunde.

Vergleich mit Mammographie und Palpation

Die Sensitivität der Mammographie war für sechs Studien angegeben und lag im Mittel bei 74,1% (54,4% bis 84,9%). Bezogen auf die Einzelstudien war die Sensitivität der Mammasonographie – bezogen auf alle Befunde – durchgehend höher. Bei Frauen bis 40 Jahren lag die Sensitivität der Mammasonographie um 11,7% (3,8% bis 21,5%) (Di Nubila et al.; Foxcroft et al. und Osako et al.) und bei Frauen bis 55 Jahren um 5,9% höher (Houssami et al.) als die der Mammographie. Houssami et al., 2003 fanden bei einer in 5-Jahresschritten altersstratifizierten Auswertung eine höhere Sensitivität bei 30-45-jährigen (14,3%), aber eine schlechtere Sensitivität bei 25-30-jährigen (11,8%).

Die Spezifität war für die Mammographie in den beiden Studien von Houssami et al. jeweils um 0,4% schlechter als für die Mammasonographie.

Morrow et al. berichteten von einer fast doppelt so hohen Rate falsch-positiver Befunde für Mammasonographie (20,3%) wie für Mammographie (10,6%) bei Frauen unter 40 Jahren.

Drei Studien (Foxcroft et al., 2004; Morrow et al., 1998; Osako et al., 2007) zeigten für Mikrokalk und ein aufgrund dieses Befunds diagnostizierten intraduktalen Karzinoms im Vergleich zur Mammasonographie eine bessere Sensitivität der Mammographie.

Vargas et al. fanden bei Frauen bis 30 Jahre für die standardisierte klinische Untersuchung die gleiche Signifikanz für das Auffinden eines echten Herdbefunds wie durch Mammasonographie und empfehlen jeweils eine Abklärung durch eine Stanzbiopsie bei 2% Karzinomrate. Foxcroft et al. fanden bei Frauen unter 20 Jahren in fast 60% der Tastbefunde kein Korrelat in der Mammasonographie. Bei histologischer Sicherung wurde kein Karzinom gefunden.

4.4.4 Diskussion

a) Methodische Diskussion

Für die vorliegende Fragestellung zum Stellenwert der Mammasonographie bei symptomatischen Frauen unter 50 Jahren konnten nur retrospektive Kohortenstudien identifiziert werden. Da nur eine der Studien – in einer retrospektiven ‚Nachbefundung‘ – verblindet durchgeführt wurde, ist eine unabhängige Bewertung des Verfahrens Mammasonographie dort nicht möglich, wo eine zusätzliche Mammographie durchgeführt wurde. Eine umfassende Bewertung der Testeigenschaften, insbesondere die Angabe der Spezifität und der prädiktiven Werte war aufgrund der Angaben oder der vorselektierten Studienpopulation (retrospektiv analysierte Brustkrebspatientinnen) nur unzureichend möglich.

b) Inhaltliche Diskussion

Im Vergleich von Mammasonographie und Mammographie zeigen die vorliegenden Studien durchgehend eine bessere Sensitivität für die Mammasonographie. Für Frauen bis 40 Jahren wurden durch die Mammasonographie im Mittel 11,7% mehr Karzinome diagnostiziert. Zu beachten ist die Anzahl falsch-positiver Befunde für Mammasonographien. Eine bessere Sensitivität liegt für die Mammographie für die Erkennung von in-situ-Karzinomen aufgrund von Mikroverkalkungen vor.

Die vorliegenden Studien schliessen unterschiedliche Altersklassen ein. Für die Empfehlung des primären Einsatzes der Mammasonographie kann aufgrund der Studienlage keine Altersgrenze angegeben werden.

Für eine validere Beurteilung der Mammasonographie als primäres diagnostisches Verfahren in Bezug auf das Alter sind prospektive konsekutive Studien mit ausreichend großen Altersgruppen für Frauen unter 50 Jahren erforderlich. Die Auswertung dieser Studien sollte altersstratifiziert erfolgen.

4.5 Anteil maligner Befunde bei mammographischer BI-RADS Klasse 3**4.5.1 Literaturrecherche**

Zu dieser Fragestellung wurden in DARE und Cochrane keine Ergebnisse identifiziert

Rechercheergebnisse PubMed

Pubmed (1969- 15.4.2007)	n= 214
Nach Screenen der Abstracts an die Expertengruppe weitergeleitete Volltexte	n= 31
Systematische Reviews:	n= 0
Einzelstudien	n= 31
Handsuche in Lit.-verzeichnissen	-
Ausgeschlossene Studien	-

Die Ergebnisse der Recherche (alle das Thema betreffende Einzelstudien) wurde ohne eine weitere Bewertung der Arbeitsgruppe zur Verfügung gestellt.

5. Evidenztabelle

5.1 Evidenz zur ergänzenden Mammasonographie beim Brustkrebs-Screening

5.1.1 Evidenz von Einzelstudien

Tabelle 1

Evidenz von Einzelstudien zur ergänzenden Mammasonographie beim Brustkrebs-Screening

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
Buchberger et al., 2000, Österreich	<p>Erhebung: prospektive Kohortenstudie</p> <p>Population: n= 8103 (med. Alter 47,6 Jahre) alle Patientinnen mit Screening-Mammographie und Palpation o.B. mit Brustdichte BIRADS 2-4 im Zeitraum erhielten MS, davon analysiert alle positiven Sonobefunde, n= 405</p> <p>Zeitraum: 1/1996-1/2000</p> <p>Befunder: n=1 spezialisierter Radiologe</p> <p>Mammographie two view</p> <p>MS-Frequenz: 5-10 oder 5-12 Mhz</p> <p>Scoring</p> <p>1. gutartig 2. nicht bestimmbar (keine Malignitätskriterien) 3. maligne</p> <p>Benignitätskriterien:</p> <p>1. homogene Hyperechogenität 2. Ellipsenförmig und dünne, echogene Pseudokapsel 3. diskrete 2-oder 3 Lobulierung und dünne, echogene Pseudokapsel</p>	<p>Scoring (n= 405)</p> <p>1. benigne n= 113 2. nicht best. n= 167 3. maligne n= 125</p> <p><u>Histologie maligne</u></p> <p>1. n= 0 0% 2. n= 3 2% 3. n= 37 30% (35 invasiv duktal, 2 in situ, med. Gr. 9,1mm)</p> <p>PPV: 30%</p> <p>Biopsie aufgrund v. MS: 3,3% n=269</p> <p>PPV Biopsie: 13,75%</p> <p><u>Rate an Karzinomen nur durch MS diagn.:</u> 22%, / 0,41% ges.</p> <p><u>Brustdichte und Karzinomentdeckung:</u></p>	<p>Zeit für Sono: 10-15 min.</p> <p>Von n= 37 Pat. mit durch MS entdecktem Karzinom hatten 54% Brustkrebs in der Anamnese</p>	<p>Zusätzlicher MS bei Frauen mit dichten Mammae (Grad 2-4) empfohlen</p>	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Malignitäts- kriterien (mind. 1): 1. nadelförmige Ausläufer (mind. 1) 2. ‚dicker‘ echogener Halo 3. Mikrolobulierung (mehr als 3 oberflächliche Lobuli) 4. gezackte Tumorränder 5. Hypoechogene(r) Bereich(e) 6. Dorsaler Schallschatten 7. erweiterter Drüsengang 8. Verästelungen Follow-up: -	Grad 2: 0,4% (2/979) Grad 3: 0,3% (24/6916) Grad 4 1,1% (12/1075)			
Corsetti et al., 2006, Italien	Erhebung: prospektive Kohortenstudie Population: n= 6449 von 17883 im Zeitraum mit Brustdichte BIRADS 3 u. 4 , Screening-Mammographie o.B. 25-96 Jahre med. Alter nicht angegeben Zeitraum: 1/2000-9/2004 Befunder: - Mammographie - ;von 4 internen und 1 externen Reviewer nachbefundet bei positivem MS zur Bestätigung des negativen Befunds MS-Frequenz: 5-10 Mhz Scoring n. angegeben Benignitäts- kriterien: - Malignitäts- kriterien: - Follow-up: -	PPV MS: nicht bestimmbar <u>1. Rate an Karzinomen nur durch MS diagnostiziert</u> 17,3% / 0,44% ges. (29/6449) <u>2. nach retrospektivem Review- Verfahren von 25 der 29 Mammographien</u> 8,9%/ 0,23% ges (15/6449) 14xpT1, 1xpT2 d.h. 10 Karzinome waren bei der Nachbefundung auf den Mammographien sichtbar Biopsie aufgrund v. MS: 7,5% (n= 486) PPV Biopsie: 3,09%	Keine Angabe des MS- Scoring-Systems Auch Kosten- berechnung: Kosten MS und Folgediagnostik gesamt: 26.7831 € Kosten pro ent- decktem Karzinom: 17.855 €	MS in dichten Mammae (Grad 3 und 4) kann hilfreich sein, die Evidenzlage lässt aber eine Empfehlung für die Routine noch nicht zu	3b
Crystal et al., 2003, Israel	Erhebung: prospektive Kohortenstudie Population: n= 1517 im Zeitraum mit Brustdichte BIRADS 2,3 und 4 und Screening-Mammographie sowie Palpation o.B. , med. Alter 52,1 SD 8,1 J. Zeitraum: 1/2000-1/2002	Scoring: 1. normal n= 1427 2. benigne n= 71 3. nicht best. n= 14 4. maligne n= 5 Histologie maligne:	Von n= 7 mit durch MS entdecktem Karzinom hatten 57% (n= 4)entweder Brustkrebs in der Anamnese oder waren High risk aufgrund der	MS ist hilfreich bei Pat. mit dichter Brust (Grad 2-4)	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Befunder: n=- ,erfahrene Befunder 2-5 J. Mammographie - MS-Frequenz: 5-12 Mhz Scoring 1. normal (o.B., simple Zyste, Duktectasie) 2. Wahrscheinlich gutartig (komplexe Zyste oder mit benignen Kriterien) 3. nicht bestimmbar 4. Maligne Benignitäts-/ Malignitäts-kriterien: s. Buchberger et al. (mind. 2 für Malignität) Follow-up: -	1. - 2. - 3. 14,3% (2/14) 4. 100% (5/5) (Gr. 4-12 mm, med. 9,6mm, alle invasiv) PPV: 33% (7/21) Rate an Karzinomen nur durch MS diagn: n.angeg. / 0,46% ges. Biopsie aufgrund v. MS: 2,5% n=38 PPV Biopsie: 13% Brustdichte und Karzinomentdeckung: Grad 2: 0% Grad 3: 0,4% Grad 4: 0,1%	Familienanamnese		
Kaplan et al., 2001, USA	Erhebung: prospektive Kohortenstudie Population: n= 1862 im Zeitraum mit Brustdichte BIRADS 3 und 4 und Screening-Mammographie sowie Palpation im Quadrant der MS-Läsion o.B. 35-87 J., med. Alter n. angegeb. Zeitraum: 7/1998-4/2000 Befunder: n=- ,technologist' oder Radiologe, erfahrene Befunder Mammographie - MS-Frequenz: 7-12 Mhz Scoring 1. negativ 2. positiv	Scoring: 1. negativ: 86,6% (1612) 2. positiv: 13,4% (250) Histologie maligne: 1. - 2. 2% (5/250) (med. Gr. 0,9cm, 0,6-14mm) alle invasiv PPV: 2% Rate an Karzinomen nur durch MS diagn.: -/ 0,26% ges.	Med. Zeit für Sono 10 min (7-20) Werte stimmen mit Studie nicht überein, da dort pro Karzinom und nicht pro Patientin gerechnet wurde Von n= 5 mit durch MS entdecktem Karzinom war 1 Pat. (20%)High risk aufgrund der Familienanamnese	MS ist hilfreich bei Frauen mit dichter Brust (Grad 3 + 4)	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Benignitäts-/ Malignitätskriterien: nicht detailliert angegeben Kriterium für Positivität: - Zyste >1cm - solider Tumor - Architekturstörung - ‚acoustic shadowing‘ Follow-up: z.T. für benigne MS-Befunde	Biopsie aufgrund v. MS: 2,7% (n= 51) 6 lost PPV Biopsie: 9,8% <u>Brustdichte und Karzinomentdeckung:</u> n. angegeb.			
Kolb et al., 2002, USA	Erhebung: prospektive Kohortenstudie Population: n= 5.418 (13547 Untersuchungen) alle asymptomatische Frauen im Zeitraum mit Brustdichte BIRADS 2, 3 und 4 von n= 11130 (27825 Untersuchungen) auch mit auffälliger Mammographie oder auff. klin. Untersuchung med. Alter 59,6 SD 15,8 Zeitraum: 1/1995 – 9/2000 Befunder: n=1 , erfahren und spezialisiert auf Brustuntersuchungen Mammographie - MS-Frequenz: 5-10 oder 5-12 Mhz Scoring 1. ohne Befund 2. benigne (Kontrolle oder normales Screening-Intervall) 3. suspekt (Biopsie) Benignitäts-/ Malignitätskriterien: nicht angegeben Verblindet: nein Follow-up: für benigne MS-Befunde	Scoring: 1. o.B 94,1% n= 12748 2. benigne 3,2% n= 441 3. suspekt 2,7% n= 358 <u>Histologie maligne:</u> 2. - 3. 10,3% (37/358) (med. Gr. 0,99 SD 6,79) PPV: 10,3% <u>Rate an Karzinomen nur durch MS diagn.:</u> 15% / 0,23% ges. Biopsie aufgrund v. MS: 2,7% PPV Biopsie: 10,3% <u>12 Mo. Follow-up für benigne Befunde zeigt kein. Karzinom.</u> <u>kein Follow-up für o.B. Befunde</u> <u>Brustdichte und Karzinomentdeckung:</u> Grad 2: 0,11% Grad 3: 0,27% Grad 4: 0,25%	Med. Zeit für MS: 4,5 min (2-11,5) Von n= 30 mit durch MS diagn. Karzinome (37 Karzinome) waren 50% Pat. mit High risk aufgrund der Familienanamnese oder einer Brustkrebsanamnese	Zusätzlicher MS erhöht die Karzinomentdeckungsrate bei Frauen mit dichter Brust (Grad 2-4)	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
Leconte et al., 2003, Belgien	<p>Erhebung: prospektive Kohortenstudie</p> <p>Population: n= 4236 von 5376 im Zeitraum, bei denen Mammographie und MS in der Einrichtung durchgeführt wurde med. Alter 60,7 J. (41-87) Versch. Ausschlusskrit: SS, Stillzeit, nicht beide Untersuchungen im Haus 73% Screening 24% Follow-up 3% palp. Befund (ausgeschlossen wurden MS-Befunde, die auch palp. waren)</p> <p>Zeitraum: 4/2000-3/2001</p> <p>Befunder: n=2 , erfahren (4 und 10 Jahre) und spezialisiert auf Brustuntersuchungen</p> <p>Mammographie two view</p> <p>MS-Frequenz: 7,5Mhz</p> <p>Scoring 1. o.B. 2. typische Zyste 3. alle anderen Befunde</p> <p>Benignitäts-/ Malignitäts-kriterien: benigne: Typische Zyste, komplett hypoechoogen mit zarter Wand und dorsaler Schallverstärkung Suspekt: alle anderen Herdbefunde oder Architekturstörung oder ‚acoustic shadowing‘</p> <p>Verblindet: nein</p> <p>Follow-up: -</p>	<p>Scoring: n. angegeben.</p> <p>Histologie maligne: n= 16 med. Gr. 1,09 cm (2-30)</p> <p>PPV: nicht bestimmbar</p> <p>Rate an Karzinomen nur durch MS diagn.: 34% / 0,37% ges.</p> <p>Biopsie aufgrund v. MS: n.angegeb.</p> <p>PPV Biopsie: nicht bestimmbar</p> <p>Brustdichte und Karzinomentdeckung: Grad 1 + 2: n= 5 Grad 3 + 4: n= 11</p> <p>Objektivität der Brustdichte-Einteilung (Interobserver-Agree): K= 0,91</p>	<p>Ergebnisse nicht nach Pat. Gruppen (Screening/Follow-up/ palp. Befund) differenzierbar!</p> <p>Zeigen, dass Mammographie bei Grad 3 + 4 Befunden hinsichtlich der Karzinomentdeckung signifikant schlechter ist als bei Grad 1 + 2 Befunden</p>	<p>MS ist hilfreich besonders zur Diagnostik bei dichter Brust (Grad 3 + 4)</p>	<p>3b</p>

5.1.2 Ausgeschlossene Einzelstudien

Tabelle 2

Ausgeschlossene Einzelstudien zur ergänzenden Mammasonographie beim Brustkrebs-Screening

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Geller et al., 2005, USA	Retrospektive Kohortenstudie Analyse von zusätzlicher Bildgebung im Rahmen des Screening – zusätzl. Mammographische bzw. sonographische Untersuchungen nicht differenziert. Ergebnis: zusätzl., Bildgebung reduziert falsch positive Befunde und erhöht die Anzahl der falsch negativen	keine Auswertung von MS alleine möglich
O'Driscoll et al., 2001, UK	prospektive Kohortenstudie bezüglich der Wertigkeit von MS bei Frauen mit mäßig erhöhtem Brustkrebsrisiko aufgrund ihrer Familienanamnese (n= 149)	spezielle Population, kleines n
Ohlinger et al., 2006, BRD	Prospektive Kohortenstudie zum Stellenwert der Mammasonographie als erstes diagnostisches Verfahren bei asymptomatischen Patientinnen (n= 448) Ergebnis: durch MS werden drei T1-Tumore diagnostiziert, die jedoch auch durch die nachfolgende MM diagnostiziert werden	Keine Untersuchung des zusätzlichen Nutzens der Mammasonographie nach Screening-Mammographie

5.1.3 Ausgeschlossenen Reviews

Tabelle 3

Ausgeschlossene Reviews zur ergänzenden Mammasonographie beim Brustkrebs-Screening

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Berg, 2004, USA	Review zum ergänzenden Einsatz von Mammasonographie beim Brustkrebs-Screening für Frauen mit dichtem Drüsenparenchym	Keine Suchstrategie angegeben, aber alle relevanten Studien bis 2004 werden diskutiert
Delorme, 2001, BRD	Review zum Einsatz der Mammasonographie und der Magnetresonanztomographie als ergänzende Verfahren beim Mammographie-Screening	Narrativ, Fokus in Bezug auf Mammasonographie liegt auf Studien, die Kriterien für Malignität untersuchten
Elmore et al., 2005, USA	Review zu verschiedenen Techniken des Mammographie-Screenings u.a. digitale Mammographie, Einsatz der Mammasonographie. Systematische Literaturrecherche	Studien von Frauen mit familiär bedingtem erhöhtem Brustkrebsrisiko und von Frauen mit mammographisch dichter Brust werden gemeinsam diskutiert, keine genaue Darstellung der jeweiligen Ergebnisse
Gordon, 2002, USA	Review zu Mammasonographie als bildgebendes Verfahren für Brustkrebs-Screening	Narrativ, aber alle relevanten Studien bis 2002 werden diskutiert
Mehta, 2003, USA	Review zum gesamten Spektrum der Indikationsstellung der Mammasonographie	Narrativ
Smith et al., 2004, USA	Review zu bildgebenden Verfahren für Brustkrebs-Screening	Narrativ, nur kleiner Abschnitt zu Mammasonographie, Fokus auf Magnetresonanztomographie
Zonderland, 2000, Niederlande	Review zum gesamten Spektrum der Indikationsstellung der Mammasonographie	Narrativ, ein Abschnitt zu Screening

5.2 Evidenz zur diagnostischen Sicherheit der Axillasonographie bei negativem klinischem Befund

5.2.1 Eingeschlossener Review

Tabelle 4
Systematischer Review zur Axillasonographie

Autor, Jahr	Alvarez, et al., 2006						
Fragestellung	Evaluierung der Sicherheit der Axillasonographie und sonographisch gesteuerter Biopsie in der präoperativen Diagnostik karzinomatös befallener axillärer Lymphknoten bei Patientinnen mit Brustkrebs						
Datenbanken/ Suchstrategie	Medline, 1980-3/2004; Suchbegriffe: 'ultrasound OR sonography AND axillary'						
Einschlusskriterien	<ul style="list-style-type: none"> ▪ präoperative Axillasonographie mit 7Mhz oder höherer Frequenz ▪ Kriterien für maligne Lymphknoten sind angegeben: Größe, Morphologie oder durch Biopsie gesichert ▪ histologisches Ergebnis der Axilladisektion oder der SNB (Goldstandard) liegt vor ▪ die Studie muss folgende Angaben enthalten: richtig und falsch positive, richtig und falsch negative 						
Ausgewählte Studien	<p>n=16 (1986-2003), 14 prospektiv, 2 retrospektiv, 8 mit konsekutiver Patientinnenselektion, 11 mit Axilladisektion als Goldstandard, 5 mit SNB; zum Teil + FNAC oder CNB</p> <p>Malignitätskriterien:</p> <table style="margin-left: 20px; border: none;"> <tr> <td>Größe > 5mm</td> <td style="text-align: right;">3x</td> </tr> <tr> <td>darstellbarer LK:</td> <td style="text-align: right;">1x</td> </tr> <tr> <td>morphologische oder strukturelle Kriterien:</td> <td style="text-align: right;">12x</td> </tr> </table> <p>Palpable und nicht-palpable LK eingeschlossen: 10x Nur nicht-palpable LK eingeschlossen: 6x</p>	Größe > 5mm	3x	darstellbarer LK:	1x	morphologische oder strukturelle Kriterien:	12x
Größe > 5mm	3x						
darstellbarer LK:	1x						
morphologische oder strukturelle Kriterien:	12x						
Ergebnisse	<p>Präoperative Axillasonographie <u>nicht-palpabler</u> LK hat bei</p> <p>a) Größe als Malignitätskriterium eine Sensitivität zwischen 48,8% (95% KI 39,6-58) und 87,1% (KI 76,1- 94,3) und eine Spezifität zwischen 55,6% (KI 44,7-66,3) und 97,3% (KI 86,1-99,9)</p> <p>b) bei Morphologie als Malignitätskriterium eine Sensitivität zwischen 26,4% (KI 15,3-40,3) und 75,9% (KI 56,4-89,7) und eine Spezifität zwischen 88,4% (KI 82,1-93,1) und 98,1% (KI 90,1-99,9)</p>						
Schlussfolgerung der Autoren	Negative Ultraschallergebnisse schließen axilläre Metastasen nicht aus						
Evidenzgrad	3a						

5.2.2 Evidenz von Einzelstudien

Anmerkung:

Im Folgenden sind die Malignitätskriterien für dargestellte Lymphknoten angegeben. Als Benignitätskriterium wurden grundsätzlich gesehen: ovale Form, symmetrisch echoarmer Kortex und echoreicher Hilus.

Tabelle 5

Evidenz von Einzelstudien zur Axillasonographie

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
Bedrosian et al., 2003, USA	Erhebung: retrospektiv	PPV: 0,5	Exzisionsbiopsie vor Sono erschien als Confounder! Zum Teil Ärzte in Ausbildung	Kombination von Sono und FNAC wird zur Reduktion von SNB empfohlen	3b
	Population: n= 208 von 513 konsekutiven mit klinisch negativer Axilla (212 hatten keine Sono, 93 hatten präop Chemo) ; med. Alter 55,4	NPV: 0,78 Sensitivität: 26,4% Spezifität: 90,9%			
	Zeitraum: 2/1994-5/2000				
	Befunder: <i>mehrere</i>	FNAC			
	Erfahrung: <i>erfahrene und unerfahrene</i>	Sensitivität: 25%			
	AS-Frequenz: 7,5-12 Mhz	Spezifität: 100%			
	Intervention: FNAC				
	% Axillae mit dargestellt. LK: -				
Kriterium für positiven LK:	1. exzentrische kortikale Verdickung oder Septierung mit fehlendem Hilus 2. hypoechogene Struktur 3. rund oder ovoid				

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	T-Stadien Tis 4% T1 77% T2 15% T3 5% T4 <1% Tx 1,5%				
Deurloo et al., 2003, Niederlande	Erhebung: prospektiv Population: n= 265, alle Patientinnen im Zeitraum für die eine SNB indiziert wurde mit klinisch negativer Axilla, 3x bilateral med. Alter 56 Jahre Zeitraum: 8/1999-1/2001 Befunder: n= 7 Erfahrung: mind. 5 Untersuchungen vor Studie AS-Frequenz: 13Mhz Intervention: FNAC % Axillae mit dargestellt. LK: 93/268 (35%) Kriterium für positiven LK: 1. atypische Kortexdarstellung 2. >/= 5mm 3. Kortex und Hilusbeurteilung nach Vasallo T-Stadien T1a-T3 (0,2-8cm) n.n. spez.	PPV: 0.74 NPV: 0.64 Sensitivität: 40,5% Spezifität: 88,4% FNAC (v.66 LK) PPV: 100 NPV: 0,73 Sensitivität: 86% Spezifität: 100%	Sehr geringe Rate an Axillae mit dargestellten LK, sehr geringe Erfahrung der Befunder, alle Fachärzte ROC-Kurven zu Malignitätskriterien zeigen, dass max. Kortex-Durchmesser und Morphologie des Kortex die verlässlichsten Malignitätsparameter sind Zeit pro Sono und FNAC 5 min.	AS + FNAC wird zur Reduktion unnötiger SLNB empfohlen Die prospektive Evaluierung der max. Kortexdicke als Malignitätskriterium wird empfohlen	2b
Esen et al., 2004, Türkei	Erhebung: prospektiv Population: n=83, unklar ob konsekutiv; klinisch negative Axilla, Zeitraum: 9/2003-7/2004 Befunder: n= 1 Erfahrung: experienced' AS-Frequenz: 7,5 Mhz, Siemens Elegra Intervention: nein aber +Dopplerflow % Axillae mit dargestellt. LK: 82/83 (99%), max. 3 beurteilt	PPV: 0.85 NPV: 0.93 Sensitivität: 94,45% Spezifität: 82,35% Werte bezogen auf 198 dargest. LK: PPV: 0.93 NPV: 0.85 Sensitivität: 80,65%	Durchschnittliche Zeit für AS: 15 min Periphere Vaskularisation ist signifikantes Malignitätskriterium, unabhängig von der Lk-Größe	Präoperativer Ultraschall wird empfohlen	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Kriterium für positiven LK: <ol style="list-style-type: none"> Kein ‚dünner‘ Kortex: Dicke > ½ Hilusdicke an breitester Stelle Asymmetr. Kortexdicke (an breitester Stelle mind. 50% mehr als an dünnster Stelle) fehlender echogen. Hilus (Kriterien nach Vasallo) periphere Vaskularisation 	Spezifität: 94,29% Dopplerflowmessung: PPV: 0.94 NPV: 0.69 Sensitivität: 50,54% Spezifität: 97,14%			
	T-Stadien	T1b bis T2 (max. 2,8cm)	Kombiniert: PPV: 0.91 NPV: 0.89 Sensitivität: 86,49 Spezifität: 93,62		
Mathijssen et al., 2006, Niederlande	Erhebung: prospektiv Population: n=131 von 174 konsekutiven Patientinnen mit klinisch negativer Axilla (75,3%) med. Alter fehlt, T1 und T2 Zeitraum: 10/99-11/2002 Befunder: n= 4 Erfahrung: - AS-Frequenz: 5-12Mhz Intervention: FNAC % Axillae mit dargestellt. LK: - Kriterium für positiven LK: <ol style="list-style-type: none"> >3mm, runde Form, hypoechoogen und heterogen T-Stadien T1 und T2 95,7%	PPV: 0.95 (0.82) NPV: 0.69 (0.7) Sensitivität: 34,0% (43,45) Spezifität: 98,7% (93,6%) FNAC-Werte nicht angegeben?	Werte in Klammern rechnen vermutlich auch pathologische Sonobefunde mit ein	Bei hoher falsch negativer Rate aber hoher Spezifität, liegt der Wert des Ultraschalls in der Vermeidung von SLN-Biopsien bei durch FNAC gesicherten pathologischem Ultraschall	3b
Nori et al., 2007, Italien	Erhebung: prospektiv Population: n=132 von 147 Patientinnen, nicht klar, ob konsekutiv; mit klinisch negativer Axilla (9 ausgeschlossen, da nicht mind. 4 LK dargestellt) Zeitraum: 7/2001-12/2003 Befunder: n= 1	PPV: 0.61 NPV: 0,77 Sensitivität: 45,2% Spezifität: 86,7% CNB (n= 14 von 31 sonographisch suspekten Befunden):	20 von 23 falsch negativen hatten Mikrometastasen, die 3 anderen waren makrometastastisch in 2, 4 und 8 LK befallen fehlender Hilus war	v.a. Mikrometastasen werden sonographisch übersehen Axillasonographie (plus CNB) kann SNB sparen	3b (+)

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Erfahrung: - AS-Frequenz: 10-13 Mhz, Esaote Intervention: CNB 14-16G ; Dopplerflow % Axillae mit dargestellt. LK: 100% (94% wenn ausgeschlossene eingerechnet) mind. 4 LK! Kriterium für positiven LK: 1. kugelig, L/T ratio =1 2. Hilum/L ratio<50% 3. fehlende Hilum-Hyperechogenität 4. exzentrische fokale Verdickung oder 'eingedellter' Kortex (denting) T-Stadien pT1a-pT2 (größter Tumordurchmesser 2,8 cm)	Sensitivität : 91,6% Spezifität : 100%	100% Malignitätskriterium		
Podkrajsek et al., 2005, Slovenien	Erhebung: retrospektiv Population: n= 165 SNB geplant, klinisch negative Axilla, med. Alter 56J Zeitraum: 1/2001-9/2003 Befunder: 3 Erfahrung: ‚experienced‘ (alle gleiches Level) AS-Frequenz: 12-15 Mhz (+Dopplerflow) Intervention: FNAC % Axillae mit dargestellt. LK: 90/165 (54%) Kriterium für positiven LK: 1. L/T-Ratio<1,5 2. Kortex max>3mm 3. fehlender Hilus 4. keine hiläre Vaskularisation T-Stadien Tis 5% T1ab 19% T1c 42% T2 33% T3 1%	PPV: 0.78 NPV: 0.77 Sensitivität: 58,5% Spezifität: 89% FNAC (von 49 suspekten LK) Sensitivität: 84% Spezifität: 91%	ROC –Analysen für L/T-Ratio (Az 0.715) - cut-off wird bei 1.2 gewählt - und max. Kortexdurchmesser (Az 0.674) 5 FNAC nicht mit repräsentativem Material	Empfehlen präoperative Sono und FNAC bei L/T-Ratio ab 1.2	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
van Rijk et al., 2006, Niederlande	Erhebung: retrospektiv	PPV: 0.53	Sehr geringe Erfahrung der Befunder, Fachärzte und Ärzte in Ausbildung Zeit pro Sono und FNAC 5 min.	Präoperativer AS + FNAC wird empfohlen, v.a. zur Vermeidung unnötiger SNB	2b
	Population: n= 726 alle Patientinnen im Zeitraum, für die eine SNB indiziert wurde mit unifokalem Karzinom und klinisch negativer Axilla; oder inflamm. Karzinom, 6 x bilateral	NPV: 0.68 Sensitivität: 34,7% Spezifität: 82,2%			
	Zeitraum: 10/99-12/2003	FNAC (v. 176 suspekten LK)			
	Befunder: n=?	PPV: 0.98			
	Erfahrung: zum Teil keine Fachärzte	NPV: 0.69			
	AS-Frequenz: 13 Mhz	Sensitivität: 61,7%			
	Intervention: FNAC	Spezifität: 99,2%			
	% Axillae mit dargestellt. LK: -	Kombiniert:			
Kriterium für positiven LK:	Sensitivität: 21%				
	1. hypoechogenes Zentrum 2. > 5mm 3. Kortex >2mm oder irregulär 4. rund oder ovoid später nur noch Kortex > 2.3mm, unklar ab wann				
T-Stadien	T1 67% T2 30% T3 22% T4 1% (inflamm.)				
Rajesh et al., 2002, UK	Erhebung: prospektiv	PPV: 0.81	Nicht beschränkt auf klinisch negative Axilla, aber Ergebnisse auch stratifiziert nach klinischem Befund Klinische Beurteilung durch einen erfahrenen Brustchirurgen	Empfehlen präoperativen AS	2b
	Population: n= 84 (2x bilateral), konsekutiv, med. Alter 60	NPV: 0.84 Sensitivität: 73,5% Spezifität: 88,5%			
	Zeitraum: -				
	Befunder: n= 1				
	Erfahrung: ‚specialist breast radiologist‘	AS bei nicht-palp. LK n= 45			
	AS-Frequenz: 7,5 Mhz	PPV: 0.5			
	Intervention: FNAC bei palpablen LK	NPV: 0.87 Sensitivität: 38%			
% Axillae mit dargestellt. LK: 100% (86/86)					

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad				
	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 15%; background-color: #e0e0e0;">Kriterium für positiven LK:</td> <td style="background-color: #e0e0e0;">1. keine ovoide Form 2. exzentrischer Kortextverbreiterung 3. fehlender echogener Hilus – ersetzt durch hypoechogenes Gewebe</td> </tr> <tr> <td>T-Stadien</td> <td>Tis 2,3% T1 26,7% T2 61,6% T3 4,7% T4 4,7%</td> </tr> </table>	Kriterium für positiven LK:	1. keine ovoide Form 2. exzentrischer Kortextverbreiterung 3. fehlender echogener Hilus – ersetzt durch hypoechogenes Gewebe	T-Stadien	Tis 2,3% T1 26,7% T2 61,6% T3 4,7% T4 4,7%	<p>Spezifität: 92%</p> <p>AS bei palpablen LK PPV: 0.88 NPV: 0.75 Sensitivität: 85% Spezifität: 80%</p> <p>Kombiniertes Assessment alle Axillae (Palpation, AS und FNAC): PPV: 1.0 NPV: 0.8 Sensitivität: 88% Spezifität: 100%</p>			
Kriterium für positiven LK:	1. keine ovoide Form 2. exzentrischer Kortextverbreiterung 3. fehlender echogener Hilus – ersetzt durch hypoechogenes Gewebe								
T-Stadien	Tis 2,3% T1 26,7% T2 61,6% T3 4,7% T4 4,7%								

5.2.3 Ausgeschlossene Studien

Tabelle 6
Ausgeschlossene Einzelstudien zur Axillasonographie

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Brancato et al., 2004, Italien	Prospektiv ; n= 155 konsekutive Fälle PPV: 0.79 NPV: 0.76 Sensitivität: 64,3% Spezifität: 86,5%	Nicht nur klinisch negative Axilla
Ciatto et al., 2006, Italien	AS und FNAC präoperativ	Nur abstract verfügbar
Couto et al., 2004, Portugal	Prospektiv AS und FNAC präoperativ n=55 bei klinisch negativer Axilla PPV: 0.60 NPV: 0.81 Sensitivität: 71,4% Spezifität: 71,4%	Zu viele Angaben fehlen, einschließlich Frequenz des Schallkopfs
Damera et al., 2003, UK	Prospektiv, n= 166 konsekutiv; PPV: 0.45 NPV: 0.71 Sensitivität: 72% Spezifität: 44,1%	Nicht nur klinisch negative Axilla
Davis et al., 2006, USA	AS + FNAC bei klinisch negativer Axilla und definiertem hohen Risiko für axilläre Metastasen	n= 27, davon viele Tumoren >5cm, viele Angaben fehlen (Sonofrequenz, Anzahl und Erfahrung der Befunder), sehr spezielle Einteilung für Metastasierungsrisiko
De Kanter et al., 2000, NL	Fokussiert auf Stellenwert ultraschallgesteuerter FNAC zum präoperativen Axillaassessment	Keine verwertbaren Angaben zu AS alleine .
Duchesne et al., 2005	Bewertet nur pathologische LK und FNAC	Keine verwertbaren Angaben zu AS alleine
Hashimoto et al., 2000	Präoperative AS-Beurteilung des den Brusttumor umgebenden Halo, falls > 0.42 cm ² als positiv für axilläre Metastasen gewertet, n= 187, retrospektiv, T1 und T2 Ergebnisse: PPV: 0.74 NPV: 0.78 Sensitivität: 69,1% Spezifität: 85,8% bei T1: NPV 0.87 Spezifität 83,1% Bei <1cm Tumorgroße Spezifität: 100%	Keine direkte Beurteilung von Lymphknoten
Kebudi et al., 2005, Türkei	Prospektiv, n = 42 PPV: 0.83 NPV: 0.73 Sensitivität: 79,1% Spezifität: 77,7%	Keine Angabe zur Palpation der Axilla
Khakpour et al., 2006, USA	Retrospektiv; 66 von 119 Patientinnen mit DCIS in der Biopsie, PPV: 0.27 NPV: 0.93	Keine Angaben zur Palpation der Axilla

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
	Sensitivität: 67% Spezifität: 72%	
Krishnamurty et al., 2002	Bewertung von FNAC bei suspekten LK	Keine verwertbaren Angaben zu AS alleine
Kuenen-Boumeester, Menke-Pluymers et al., 2003, Niederlande	Fokussiert auf Auswertung von AS-gest. FNAC zum präoperativen Axillaassessment	Keine verwertbaren Angaben zu AS alleine
Lemos et al., 2005	AS, Dopplerflow und FNAC präoperativ	n= 40, viele Angaben fehlen : Frequenz des Schallkopfs, T-Stadienverteilung, Anzahl und Erfahrung der Befunder, genaue Definition der Kriterien eines suspekten LK
Lumachi et al., 2006, Italien	Prospektiv ;n=77 konsekutiv PPV: 0.76 NPV: 0.73 Sensitivität: 67,6% Spezifität: 80%	Nicht nur klinisch negative Axilla
Motomura et al., 2001, Japan	AS wird nach Darstellung eines Hotspots (Gamma-Probe) ausgeführt	Keine Bewertung von AS alleine möglich
Nori et al., 2005, Italien	Studie ist Teil einer größeren Auswertung, die 2007 publiziert wurde	Daten werden in der Publikation von 2007 (größeres Kollektiv) auch ausgewertet
Ohta et al., 2000	Vergleich der Efficacy zur LK-Beurteilung von AS und PET	n= 30, viele Angaben fehlen (T-Stadien, Befunder, genaue Charakterisierung von suspekten LK), 66% axilläre Metastasen im Kollektiv, dient nicht der Früherkennung
Oruwari et al., 2002, USA	Fokussiert auf AS und FNAC bei lokal fortgeschrittenem Mammakarzinom >3cm, n= 27, 70% LK-Metastasen	Dient nicht der Früherkennung
Popli et al., 2006, Indien	AS und FNAC präoperativ	n= 30, viele Angaben fehlen, Frequenz des Schallkopfs, Anzahl und Erfahrung der Befunder, LK palpabel oder nicht
Sapino et al., 2003, Italien	Prospektiv, n=267, konsekutiv, 1/2000-12/2000 PPV: 0.63 NPV: 0.84 Sensitivität: 68,2% Spezifität: 80,4%	Nicht nur klinisch negative Axilla
Sato et al., 2004, Japan	Prospektiv, n= 262, konsekutiv PPV: 0.93 NPV: 0,7 Sensitivität: 44,6% Spezifität: 97,3%	Keine Angaben zur Palpation, zum Teil massiver Axillabefall
Somasundar, et al., 2006, USA	Fokussiert auf Auswertung von AS-gest. FNAC zum präoperativen Axillaassessment	Keine verwertbaren Angaben zu AS alleine
Topal et al., 2005, Türkei	Bewertet nur pathologische LK und CNB	Keine verwertbaren Angaben zu AS alleine
Yang et al., 2000, Singapur/ Hongkong	Präop. AS + Dopplerflow Evaluation bei palpablen und nicht palpablen LK	Ergebnisse sind pro LK dargestellt, nicht pro Patient
Zgajnar et al.,	Fokussiert auf Risiko von Macrometastaen in Nicht-	n= 36 von 56 (20 Patientinnen)

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
2005, Slovenien	Sentinel LK bei negativem AS und Micrometastasen im Sentinel	lehnten Axilladisektion nach SNB ab), Risiko von Bias, spezielle Fragestellung
Zgajnar et al., 2006, Slovenien	Retrospektive Fall-Kontrollstudie	Viele Angaben fehlen

Tabelle 7

Ausgeschlossene Reviews zur Axillasonographie

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Balu-Maestro et al., 2005, Frankreich	Darstellung der Vor- und Nachteile bildgebender Verfahren zur Beurteilung der Axilla bei Brustkrebs	Narrativ
Lernevall 2000, Dänemark	Darstellung der Vor- und Nachteile bildgebender Verfahren zur Beurteilung axillärer Lymphknoten	Narrativ

5.3 Mammasonographie bei auffälligem Brustbefund

5.3.1 Evidenz von Einzelstudien

Tabelle 8

Evidenz von Einzelstudien zur Mammasonographie bei auffälligem Brustbefund

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
Berg W.et al., 2004	<p>Erhebung: prospektive Kohortenstudie zur Evaluierung der Sicherheit von Klin. Untersuchung, MM, MS und MRT</p> <p>Population: n= 111 Pat , 120 Mammae mit 258 Befunden, davon 80 (73%) mit palpablem Befund; med. Alter: 48J, (26-81)</p> <p>Zeitraum: 9/1999-2/2002</p> <p>Befunder: 2 breast radiologist</p> <p>Erfahrung: 2 und 10 Jahre</p> <p>MS-Frequenz: 7,5 Mhz und 10Mhz</p> <p>Scoring-System angegeben: 1. simple Zyste 2. wahrscheinlich benigne 3. suspekt 4. hochwahrscheinlich maligne</p> <p>Referenzstandard: histologische Beurteilung:</p> <p>Beurteilung verblindet: nein, nur MS zu MRT</p>	<p><u>Karzinome gesamt:</u> 177</p> <p><u>MS:</u> Sensitivität: 83% (147/177) Spezifität: 34% (28/81)</p> <p>MS entdeckt 18 invasive Karzinome und 9 DCIS, die mammograpisch okkult waren, ins. n= 27 (15,23%)</p> <p>Mediane Größe invasiver Tumoren 27 mm (4-90)</p> <p><u>MM:</u> Sensitivität: 67,8% (120/177) Spezifität: 75% (61/81)</p> <p><u>MRT:</u> Sensitivität: 94,4% (167/177) Spezifität: 26% (21/81)</p> <p>Klin Untersuchung: Sensitivität: 50,3% (89/177) Spezifität: 92% (75/81)</p> <p>Beste Kombination: Klin U. + MM + MRT</p>	<p>Ergebnisse in Bezug auf palp. Befunde nicht gesondert ausgewiesen</p> <p>Sensitivität Mammographie war invers korreliert zur Brustdichte (p<0.01)</p>	<p>Beste Kombination: Klin U. + MM + MRT, aber Klin. U, MM und MS fast genauso effektiv</p> <p>(MRT hat v.a. höhere Sensitivität für DCIS)</p>	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
Cid et al., 2004, UK	Erhebung: prospektive Kohortenstudie zur Evaluierung von MS bei Diskrepanz Pat.-bef. und ärztl. Befund	MS: o.B. 83% (n= 350) mit Befund: 17% (n= 70) Zysten: 69% (n= 48) Solide Bef. 31% (n= 22)	Standardisierte klinische Untersuchung beschrieben, ärztl Untersuchung ist negativ	MS ist eine wichtige Ergänzung zu negativer klinischer Palpation bei durch die Pat. gefühlten Befund	2b
	Population: n= 420 Pat. mit selbst getastetem Knoten, der bei standardisierter klinischer Untersuchung nicht nachvollziehbar war med. Alter: MS o.B: 36,6J.(14-69) MS m.B: 42,2J (18-66) P<0.001.	5 Befunde klass. als nicht-bestimmbar oder verdächtig, davon 3 Karzinome (alle <7mm), 1 okkult in der Mammographie: MS allein entdeckt 33% der Karzinome (n= 1 <1cm)			
	Zeitraum: 7/1999-4/2001	NPV: MS: 0.98			
	Befunder: 1 breast radiologist				
	Erfahrung: n.n. spez.				
	MS-Frequenz: 7,5 Mhz und 13Mhz				
	Scoring-System angegeben: nein				
	Referenzstandard: histologische oder zytologische Beurteilung, Verlaufsbeobachtung med. 3,7 Jahre				
	Beurteilung verblindet: nein				
Flobbe et al., 2003, Niederlande	Erhebung: prospektive Kohortenstudie an einem Zentrum	Sensitivität: Klin. U.: 9,0% MM: 82,9% MS : 87.6% Spezifität: Klin. U: 92,5% MM: 91,9% US: 95,5%	(Werte beziehen sich auf Untersuchungen pro Brust, nicht auf Anzahl der Patientinnen)	Empfehlen MS grundsätzlich	2b
Population: n= 470 Pat. mit palp. Befund von 2720 konsekutiven mit MS im Zeitraum med. Alter ges: 50,2J (25-78)					
Zeitraum: 10/1999-8/2000					
Befunder: -					
Erfahrung: alle mit mehr als 5 Jahren Erfahrung					
MS-Frequenz: 12,5 Mhz					
Scoring-System angegeben: nein					

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Referenzstandard: histologische Beurteilung, Verlaufsbeobachtung (mammogr. und klinisch) mind. 14 Monate Beurteilung verblindet: nein	PPV MM + Klin. U.: 52/185 28,1% PPV US+MM+Klin U.: (125/319) 39,2% im Gesamtkollektiv durch MSalleine werden 8 zusätzl. Karzinome entdeckt (6%), keine Größe angegeben MSstuf 132 durch MM und klin. Untersuchung als positiv eingeschätzte Befunde richtig zurück FP US: n= 46 FN US: n= 1 <u>Altersspezifische Analyse</u> <u>Aerea under the ROC-Curve : Zugewinn durch US:</u> <u><50J: 0,09</u> <u>(von 0,87 auf 0,98)</u> <u>>50J: 0,05</u> <u>(von 0,95 auf 0,99)</u>			
Georgian-Smith et al., 2000, USA	Erhebung: retrospektive Subgruppenanalyse eines Multicenter-Trials (Kohortenstudie) Evaluierung der FN im MS bei palp. Befund Population: n= 293 (alle malignen) von 616 prospektiv eingeschloss. Pat. mit palpablem Befund (med. Alter 46J, 16-95) aus 14 akademischen Institutionen Zeitraum: 1/1992-3/1994	FN: Falsch negative MS: 0% Mammographie: 6,1% (n= 18, med. Gr: 1,7 cm, 14x invasiv.-duktales Karzinom) PPV, NPV können aus den Daten nicht berechnet werden	Med. Größe alle: 2,1 cm Studie gesponsert von LTA	Prospektive Studie erforderlich, die ein Kollektiv aller Frauen einschließt, auch mit klinisch u. sonographisch benigne erscheinenden Befunden mit entsprechender Verlaufskontrolle um FN sicher zu beurteilen	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Zeitraum: 9/1999-5/2000 Befunder: n= Erfahrung: alle erfahren MS-Frequenz: mind. 7,5Mhz Scoring-System angegeben: 1. hochwahrscheinlich benigne 2. nicht bestimmbar 3. maligne (nach sonomorphologischen Kriterien von Stavros) Beurteilung verblindet: nein	Karzinome n= 19 suspekta Befunde (31,6%) nur im MS darstellbar Davon maligne: n= 2 (d.h. MS entdeckt 14,2% der Karzinome) keine Größenangabe FP MS: nicht angebbbar, da nicht alle mit Biopsie <u>NPV MS+MM:</u> 100%			

Ergänzende Studien von 1999

Moss et al., 1999, UK	Erhebung:	prospektive Kohortenstudie an einem Zentrum	MS ges:	Pat. < 35 J Sono als erste Untersuchung und > 35 MM als erste MS	Empfehlen grundsätzlich MS als zusätzliche bildgebende Untersuchung	3b
	Population:	n= 559 mit palpabler Auffälligkeit, davon in 66% MM+MS, restl. nur 1 Untersuchung med. Alter nicht angegeben	Sensitivität: 88,9%			
	Zeitraum:	1/1993-7/1996	Spezifität: 77,9%			
	Befunder:	n= 4	MM ges:			
	Erfahrung:	alle erfahren	Sensitivität: 78,9%			
	MS-Frequenz:	n. angegeben	Spezifität: 82,7%			
	Scoring-System angegeben:	1. o.B. 2. benigne 3. möglicherweise maligne 4. wahrscheinlich maligne	Kombiniert:			
Referenzstandard:	histologische Beurteilung	Sensitivität: 94,2%				
		Spezifität: 67,9%				
		Befunde ins:				

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Beurteilung verblindet:	1. Beurteilung verblindet, 2. definitive Beurteilung in Kenntnis beider Befunde	303 benigne 256 maligne <u>MS</u> bei MM normal oder benigne: n= 170 zusätzl. diagnostizierte Karzinome: n= 30 (11,7%) keine Größenangabe der Karzinome FN = 12 (als benigne klassifiziert) FP = 22 1x MS+MM beide falsch normal			
Rahbar et al., 1999, USA	Erhebung:	retrospektive Kohortenstudie an einem Zentrum	n= 38 maligne Befunde, für 32 MM vorhanden	MM vorhanden für n= 133 gesamt (77,3%)	Zusatznutzen von MS bei palp. Befund und MM ist nicht statistisch signifikant	2b
Population:		n= 161 konsek. Pat. (162 Befunde , n= 118 palp. 73%) med. Alter 47 J, (22-75J)	Sensitivität: MM 89% MM + MS 95%	Wesentliches Ziel der Untersuchung war die Evaluation von MS- Malignitäts und Benignitätskriterien und die Interobserver- variabilität		
Zeitraum:		1/1994-171996	P= 0.07			
Befunder:		retrospektiv n= 3	Keine Größenangabe der zusätzl. diagn. Karzinome			
Erfahrung:		alle erfahren in Brustdiagnostik				
MS-Frequenz:		7 und 10 Mhz				
Scoring-System angegeben:		ja, morphologische und Score nach BIRADS resultierend in benigne (1-3) und maligne (4 u. 5)	Spezifität: MM 42% Mm + MS 42% n.s.			
Referenzstandard:		histologische Beurteilung durch CNB (140) oder zytologische Beurteilung durch FNAC (22)	PPV: MM 32% MM + MS 34% n.s.			
Beurteilung verblindet:		ja, retrospektiv	NPV: MM 92% MM+MS 96% n.s.			

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
Skaane 1999, Norwegen	Erhebung: prospektive Kohortenstudie an einem Zentrum	Durch MS werden zusätzlich 11 Karzinome bei falsch negativer MM entdeckt (4,8%)	Weitere Werte können aus den Angaben nicht berechnet werden	Empfehlen MS in Kenntnis des MM-Befundes v.a. bei palpablem Befund und benigner oder nicht bestimmbarer MS-Diagnose	3b
	Population: n= 228 Pat. mit palp. Befunden (70%) von 327 konsekutiven Mammakarzinomen med. Alter 64 Jahre	MS diagnostiziert 20 (42%, 8,8% gesamt) von 48 palpablen Tumoren mit benigner oder nicht bestimmbarer MM Diagnose richtig als maligne			
	Zeitraum: 1/1988- 1/1992	Durch MS zusätzl zur MM richtig diagnost. Karzinome gesamt: 13,6%+			
	Befunder: n= -				
	Erfahrung: alle erfahren in Brustdiagnostik				
	MS-Frequenz: 7,5 Mhz	Durch MM werden 41 Karzinome bei falsch negativem Ultraschall entdeckt (18%)			
	Scoring-System angegeben: normal/ Zyste/ undefinierbarer Herdbefund/ fokale Auffälligkeit/ Tumor – eingeteilt nach benigne/unbestimmbar/maligne	Keine Größenangabe der Karzinome			
	Referenzstandard: Zytologische oder histologische Beurteilung oder 12 Monate mammographische Kontrolle				
Beurteilung verblindet: nein					
Skaane et al., 1999, Norwegen	Erhebung: retrospektive Kohortenstudie an einem Zentrum	ROC-Kurven-Analyse ges.: MS: A= 0.846	MM nicht konsekutiv ausgewählt Keine absoluten Zahlenangaben verfügbar	MS im Vergleich zur MM hat im vorliegenden Kollektiv nur eine höhere Sensitivität für Tumoren > 2cm, die Autoren empfehlen bei unklarem MM Befund grundsätzlich eine MS+MM kombinierte Beurteilung	3b
	Population: n= 115 von 200 retrospektiv ausgewählten Pat. mit palp. Befunden, mit MM-Befund ohne Mikroverkalkungen, nicht eindeutig maligne in Erstbefundung	MM: A= 0.831 n.s. MS+MM: A= 0.919 P< 0.05			
	Zeitraum: n. angegeben	ROC-Kurven-Analyse für Tumoren < 2cm:			
	Befunder: retrospektiv n= 4	MS: A= 0.827			
	Erfahrung: alle erfahren in Brustdiagnostik	MM: A= 0.875			
	MS-Frequenz: 7,5 Mhz				

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Scoring-System angegeben: normal/ Zyste/ undefinierbarer Herdbefund/ fokale Auffälligkeit/ Tumor – eingeteilt nach benigne/wahrscheinlich benigne/unbestimmbar/wahrscheinlich maligne/maligne Referenzstandard: Histologische Beurteilung Beurteilung verblindet: ja, retrospektiv zunächst 50% der Bilder verblindet, dann MM und MS kombinierte Beurteilung	n.s. MS+MM: A = 0.914 n.s. ROC-Kurven-Analyse für Tumoren >2cm: MS: A = 0.857 MM: A = 0.780 P < 0.01 MS+MM: A = 0.926 P < 0.001			
Zonderland et al., 1999, Niederlande	Erhebung: prospektive Kohortenstudie an einem Zentrum mit 2 Standorten (Subgruppenanalyse) Population: n= 262 Pat. mit palp. Befunden mammographisch okkult mit MS von insgesamt 4728 konsekutiven Pat. mit MM, davon 918 mit palpablen Befunden med. Alter 53 J, (30-94) Zeitraum: 1/1994-1.7.1996 Befunder: n= 5 Erfahrung: alle erfahren in Brustdiagnostik MS-Frequenz: 7,5 und 10 Mhz Scoring-System angegeben: 1-4 Referenzstandard: Zytologische oder histologische Beurteilung oder 12 Monate mammographische Kontrolle	In 918 Pat. mit palpablem Befund wurden n= 55 Karzinome gefunden, davon allein durch MS (bei n= 262 Pat.) n= 17 (30,9%). n= 4.728 (palpable und nicht-palpable Befunde) Sensitivität: MM ges: 86% Spezifität: MM ges: 97% Sensitivität MM +MS: 95% (p< 0.01) Spezifität MM+MS: 98% keine Größenangabe der durch MS alleine diagn. Karzinome	Sensitivität und Spezifität nicht für MS getrennt ausweisbar	Empfehlen grundsätzlich MS	3b

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Beurteilung verblindet:	nein	Der Zugewinn an Sensitivität betrug bei 30-39j. 22% 40-49j. 18% 50-59j. 7% 60-69j. 5% Der Zugewinn an Spezifität betrug bei 30-39j - 2% 40-49j + 2% 50-59j + 8% 60-69j + 5%			

5.3.2 Ausgeschlossene Studien

Tabelle 9

Ausgeschlossene Einzelstudien zu Mammasonographie bei auffälligem Befund

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Arger et al., 2001, USA	retrospektive Kohortenstudie n= 70 Herdbefunde beurteilt von 3 erfahrenen Befundern (verblindet zu MM und klin. Unters.), reiner Mikrokalk ausgeschlossen Sicherheit für MS-Malignitätskriterien und Benignitätskriterien, Einteilung in ein Scoring System 0-5 und Analyse der Interobserverobjektivität PPV Maligne: zwischen 55,7% und 62,8% PPV benigne: zwischen 96% und 100% Interobserverobjektivität	Anteil palpabler Befunde nicht klar Fokus auf Bewertung einzelner Kriterien für Malignität
Baez et al., 2005, BRD	Prospektive Kohortenstudie n= 398 Herdbefunde mit CNB oder FNAC, konsekutiv 8/1997-2/2001 med. Alter 49,8J, 7,5 Mhz, 3 erfahrene Befunder Scoring mit 7 Merkmalen n. Stavros, n. EUSOMA eingeteilt in 1-4. II Benigne n= 278 (69,8%) PPV f. Benignität: 99,3% (276/278) III Unklar n= 70 (17,6%) PPV f. Malignität: 22,8% (16/70) IV suspekt n= 50 (12,6%) PPV 84% (42/50) NPV und PPV für die Malignitätskriterien auch einzeln berechnet	Anteil palpabler Befunde unklar
Berg, et al., 2000, USA	Sicherheit von MS zur Diagnose von multizentrischem oder multifokalem Karzinom	Fokus liegt auf Entdecken von (nicht-palpablen) Zusatzbefunden
Beyer et al., 2003, USA	Retrospektive Kohortenstudie zum NPV bei negativer MM und MS bei palpablem Befund FN: 3,7% (n= 13 von n= 351) für MS + MM	Keine Angaben zu MS alleine
Cha et al., 2005, Korea	Prospektive Kohortenstudie zum Vergleich der Sicherheit von konventionelle Ultraschall und spatial compound Bildern n= 71 Pat., Vergleich nicht stat. signifikant unterschiedlich	63% nicht palpable Läsionen
Chala et al., 2007, Brasilien	Retrospektive Kohortenstudie solider MS-Herdbefunde im Hinblick auf Sicherheit für Benignität um Anzahl von benignen Biopsien zu reduzieren n= 203 Pat. Bei Verzicht auf Biopsie von nach Scoring wahrscheinlich benigne hätte zu einer 42% Reduktion der Biopsien geführt	Nicht angegeben, wie viele Patientinnen einen klinisch auffälligen Befund hatten, Pat. mit negativer MM wurden von der MS-Beurteilung ausgeschlossen.
Dennis et al., 2001, USA	Retrospektive Kohortenstudie zur Wertigkeit von normalem MS + MM bei palpablem Befund n= 486 Pat. mit n= 60 Biopsien und n= 435 Pat. mit MM und/oder klin. Follow-up mind 2 Jahre NPV = 100%	Keine Angaben zu MS alleine
Greene et al., 2006, USA	Retrospektive Beurteilung der Klinik und Bildgebung von 4091 MS-gesteuerten Biopsien. n= 452 Maligne Biopsiebefunde n= 32 MM okkult davon 19 palpable und 10 nicht palpabel bei 3 unklar. n= 9 nur durch MS entdeckt	MS entdeckte Befunde klinisch unauffällig
Hong et al., 2005, USA	Retrospektive PPV und NPV für MM und MS-BIRADS-Kategorien von soliden Herdbefunden; 202 palpable Befunde und 201 nicht palpable	Keine Evaluierung nach palpablen und nicht palpablen Herdbefunden

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
	<p>n= 66 (16%) MS bei Herdbefund und neg. MM</p> <p>Kriterien nach Stavros werden bewertet. Beste PPV für Malignität: ‚Spiculation‘ 91% (21/23) ‚Höher als breit‘ 71% (76/107)</p> <p>Beste NPV für Malignität: dünne echogene 95% (63/66) Pseudokapsel ‚Circumscribed margin‘ 90% (173/193)</p> <p>Werte für MM nicht dargestellt</p>	
Mitnick et al., 1999, USA	Retrospektive Kohortenstudie von 78 tubulären Karzinomen aus 932 konsekutiven Karzinomen im Zeitraum (1990-1997) zeigt PPV MM mit 87% und PPV MS (durchgeführt bei n= 38 Fällen) von 100%, davon 9x MM neg,	palpabler Befund nur in 15 Fällen (19%)
Moy et al., 2002, USA	Retrospektive Kohortenstudie von n= 829 Pat. mit palpablem Befund und neg. MM und MS, Follow-up von 233 Pat. (28,2%) mehr als 2 Jahre , davon 6 (2,6%) mit Mammakarzinom	Keine Angaben zu MS alleine , Follow-up nur von ca. 28% verfügbar (möglicher Selektionsbias)
Soo et al., 2001, USA	Retrospektive Kohortenstudie von n= 420 Pat. mit 455 palp. Befunden mit neg. MM und MS und mittl. Follow-up von 25 Monaten FN: n= 1 NPV: 99,8%	Keine Angaben zu MS alleine
Taylor et al., 2002, USA	Prospektive Kohortenstudie in 9 Zentren 1994 durchgeführt, n= 705, 51,5% palpabel, 48,5% nicht palpabel MS+MM versus MM alleine: Sensitivität gleich (97,9 versus 97,1%) Spezifität besser (63,8 versus 51,4%) PPV besser (55,3 versus 51,4%) NPV (98,3 versus 97,4)	Zu palpablen Befunden nur Diff. Area under the ROC curve MS+MM versus MM allein angegeben: (Diff. 0.290 nicht statistisch sign.) Keine weiteren Werte für palpable Befunde berechenbar

Tabelle 10

Ausgeschlossene Reviews zu Mammasonographie bei auffälligem Brustbefund

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Harvey, 2006, USA	MS palpabler Brust-Befunde	Narr. Review
Hille et al., 2004	Neubewertung der Mammasonographie in der Mammakarzinomdiagnostik	Narr. Review, viele relevante Studien werden diskutiert, aber keine Suchstrategie, keine Einschlusskriterien
Pruthi, 2001, USA	Erkennung und Bewertung palpabler Brust-Befunde	Narr. Review

5.4 Stellenwert der Mammasonographie Frauen unter 50 Jahren mit auffälligem Brustbefund

5.4.1 Evidenz von Einzelstudien

Tabelle 11

Evidenz von Einzelstudien zur Mammasonographie bei Frauen unter 50 Jahren mit auffälligem Brustbefund

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
Di Nubila, 2006, Italien	Erhebung: retrospektive Kohortenstudie an einem Zentrum	US: Sensitivität: 88,7% (212 v.239)	87% der Pat. hatten Brustdichte Grad 3 und 4 9 ausgeschl. wegen fehlendem Bildmaterial n=27 (11,3%) MM: FN: 35 (15,1%) (vor allem in dichtem Gewebe Kategorie 4) Unterschied ist statistisch nicht sign. n= 20 mit beiden Methoden falsch negativ (16 als benigne klassifiziert, 4 ohne Korrelat) 7 sichtbar mammographisch aber nicht sonographisch 15 sichtbar sonographisch aber nicht mammographisch	Empfehlen MS als erstes diagnostisches Verfahren und MM ergänzend – in Kenntnis des klinischen und MS Befunds: dadurch kann nach Meinung der Autoren die Sensitivität der MM erhöht werden (wie im vorliegenden Setting)	3b
	Population: n= 239 von n= 348 Pat. <35 Jahre mit Erstdiagn. Brustkrebs med. Alter 32J (22-35) 97,2% symptomatisch	MM: Sensitivität: 84,9% (198/233) Unterschied ist nicht statistisch signifikant			
	Zeitraum: 1/2000-12/2002	US:			
	Befunder: -	FN:			
	Erfahrung: alle mit Spez. in Brustdiagnostik				
	MS-Frequenz: 7,5 Mhz und 13Mhz				
	Scoring-System angegeben: 1 =normal 2 = benigne 3 = wahrscheinlich benigne 4 = V.a. Malignität	MM: FN: 35 (15,1%) (vor allem in dichtem Gewebe Kategorie 4)			
Referenzstandard: histologische Beurteilung	Unterschied ist statistisch nicht sign.				
Beurteilung verblindet: nein					
Foxcroft et al., 2001,	Erhebung: retrospektive Kohortenstudie an einem Zentrum	MS: 58,8% (307/523 mit Herdbefund) o.B.	Es wurde keine Mammographie	Empfehlen MS bei palpablem Befund, keine	4

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
Australien	Population: n= 619 (98%) Pat. < 20 J. (9-19, 63% 18 o.19J) mit MS von n= 634 (98%) mit Symptomen (89% Tastbefund) Zeitraum: 1/1990-12/1999 Befunder: - Erfahrung: alle mit Spez. in Brustdiagnostik MS-Frequenz: n. beschrieben Scoring-System angegeben: - Referenzstandard: histologische oder zytologische Beurteilung (n= 172) klinische und sonographische Nachbeobachtung für n= 28 mit V.a. Fibroadenom o.B. Beurteilung verblindet: Nein	FN: (n= 1/21 durch FNA untersuchte): 4,7% 195/216 positiven MS-Befunden zytologisch oder histologisch gesichert: Zyto-/Histologische Diagnose: 26,5% Fibroadenome 7,1% Zysten 0,5% benigne Phylloides tumore kein Karzinom	eingesetzt PPV und NPV kann aufgrund der Angaben nicht berechnet werden	MM. Bei benignem MS- Befund Verlaufskontrolle Vorzug vor zytologischer/ histologischer Sicherung geben	
Foxcroft et al., 2004, Australien	Erhebung: retrospektive Kohortenstudie an einem Zentrum Population: n= 239 (von 30961 ges.) Pat. <40 Jahre (20-39, 92% >30J) mit ED Brustkrebs, 87% mit palp. Befund Zeitraum: 2/1992-2/2002 Befunder: n= ? Erfahrung: alle erfahren in Brustdiagnostik MS-Frequenz: high resolution, Frequenz nicht angegeben Scoring-System angegeben: 1. kein Befund 2. benigne 3. wahrscheinlich benigne 4. V.a. Malignität Referenzstandard: histologische Sicherung (FSE) Beurteilung verblindet: nein	US: (v. 230 verfügbaren) Sensitivität: Befund: 92,2% Sensitivität: suspekter Befund: 74,5% FN: (Befund): 6.6% MM: (v. 215 verfügbaren) Sensitivität: Befund: 72% Sensitivität: suspekter Befund: 54,4% FN: (Befund): 18,3% MM zeigte suspekten Befund in 9,4% in denen MSkeinen oder einen benignen Befund zeigte (Mikrokalk) bei insgesamt weit besserer Sensitivität von US (Signifikanz nicht berechnet)	13,7% hatten eine Familienanamnese 1. Grades Brustdichte war negativ korreliert mit auffälligem mammographischen Befund 79% der Karzinome in Brustdichte Grad 3 oder 4	Empfohlen wird bei symptomatischen Frauen < 40 MS, dann Mammographie und zytologische Sicherung vor OP	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
Houssami et al., 2002 Australien Kollektiv aus Italien	Erhebung: retrospektive Kohortenstudie an einem Zentrum altersstratifiziert	MS: Sensitivität: 80,5%	Primäre Untersuchung im Kollektiv Bis 40 J., MS, dann MM Cave: Untersuchungen jeweils in Kenntnis des anderen bildgebenden Befunds	Autoren empfehlen prospektive Studien zur Bestätigung der altersabhängigen Befunde mit verblindet durchgeführter Bildgebung	3b
	Population: n=957 Pat mit Symptomen und neu diagnostiziertem Brustkrebs mit MS und MM, med. Alter 58J (24-95)	MM: Sensitivität: 77,6%			
	Zeitraum: 1986-1997	Unterschied nicht stat. signifikant			
	Befunder: 5 Radiologen				
	Erfahrung: alle ‚erfahren in Brustbeurteilung	Sensitivität: (%) Stratifiziert nach Alter: N MS MM 20-34 27 59,3 70,4 35-39 58 70,7 56,9 40-44 107 79,4 64,5 45-49 132 75,8 65,9 (weitere Altersklassen nicht aufgeführt)			
	MS-Frequenz: 10 Mhz				
	Scoring-System angegeben: 1. o.B. 2. <i>benigne</i> 3. nicht bestimmbar oder verdächtig 4. suspekt/maligne (3 und 4 wird als + gewertet)				
	Referenzstandard: histologische oder zytologische Beurteilung				
Beurteilung verblindet: nein	Sensitivität bis 62 (KI 49-67) besser mit MS, dann MM (stat. Signifikanz nicht ber.) keine Angabe zu Spezifität				
Houssami et al., 2002, Australien Kollektiv aus Australien	Erhebung: retrospektive Kohortenstudie	MS: Sensitivität: 83,8%	Primäre Untersuchung im Kollektiv bis 30 J., MS, nach 30 J. Mammogr	Autoren empfehlen prospektive Studien zur Bestätigung der altersabhängigen Befunde mit verblindet durchgeführter Bildgebung	3b
	Population: n= 240 Pat. 27-55J. (med. 44,9) mit Sympt. und hist. verifiziertem Karzinom (von ins. 258, konsekutiv) und n= 233 Kontrollen mit benignem Befund	MM: Sensitivität: 77,9%			
	Zeitraum: 1994-1996	MS nicht stat. signifikant			
	Befunder: 2 Radiologen, bei Dissens Entscheidung durch einen Dritten	Komb. MS+MM: 95,4% (1 Test pos.)			
	Erfahrung: alle drei ‚erfahren‘ in Brustbeurteilung				
MS-Frequenz: 7,5 Mhz und 11,5Mhz					

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Scoring-System angegeben:	1. o.B. 2. <i>benigne</i> 3. nicht bestimmbar 4. suspekt 5. maligne (3-5 als +)	Sensitivität: MS: 96,6% Mammographie: 97% MS + MM: 94,4%			
Referenzstandard:	histologische oder zytologische Beurteilung					
Beurteilung verblindet:	nein					
Houssami et al., 2003, Australien Gleiches Kollektiv wie oben!	Erhebung:	retrospektive Kohortenstudie Prospektive an einem Zentrum altersstratifiziert Re-Bewertung von Bildmaterial (Mammographie und MS)	MS: Sensitivität: 81,7% MM: Sensitivität: 75,8% Diff: 5,9% (-1,5-13,2) n. stat. sign. Spezifität: MS: 88% Mammographie: 87,6% Stratifiziert nach 5 J. Altersgruppen (<35J, 36-40 etc.) ergibt sich für Sensitivität und Spezifität der beiden Untersuchungen ebenfalls kein stat. sign. Unterschied. Sensitivität MS stratifiziert nach </> 45J Ergibt stat. signifikanten Unterschied von 13,2% (CI 2,1-24,3%) p=0,035 zugunsten des MS (84,9%).	n= 18 ausgeschlossen wegen fehlendem Bildmaterial Primäre Untersuchung bis 30 J. MS, nach 30 J. Mammogr. Vorstellungsgrund in 80% palpabler Befund. Retrospektive Auswertung von MS- Bildern, die in Kenntnis des mammo- graphischen Befunds angefertigt wurden 45 J. retrospektiver Cut- off	Empfehlen MS als erstes bildgebendes Verfahren bis 45 J. Zur Bestätigung wird eine prospektive Studie gefordert	2b (-)
Population:	n=240 Pat. 27-55J. (med. 44,9) mit Sympt. und hist. verifiziertem Karzinom (von ins. 258, konsekutiv) und n= 233 Kontrollen mit benignem Befund					
Zeitraum:	1994-1996					
Befunder:	2 Radiologen, bei Dissens Entscheidung durch einen Dritten					
Erfahrung:	alle drei ‚erfahren‘ in Brustbeurteilung					
MS-Frequenz:	7,5 Mhz und 11,5Mhz					
Scoring-System angegeben:	s.u.					
Referenzstandard:	histologische oder zytologische Beurteilung					
Beurteilung verblindet:	ja, bei jetziger Auswertung					

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
Malik et al., 2006, Pakistan	Erhebung: retrospektive Kohortenstudie an einem Zentrum	MS-Scoring: 1. 49 benigne 2. 5 nicht bestimmbar 3. 2 maligne		MS eignet sich als primäre diagnostische Bildgebung in dieser Altersgruppe; hat eine höhere Sensitivität für die Entdeckung benigner Befunde	3b
	Population: n= 56 Pat 15-25J mit palpablem Befund	FNA: 53 benige 3 maligne			
	Zeitraum: 2/2002-4/2005				
	Befunder: -				
	MS-Frequenz: 7,5 Mhz	MS PPV: 75%, 43% (bei 2+3) MS NPV: 100%			
	Scoring-System angegeben: 1. benigne 2. nicht bestimmbar 3. maligne (Verwendung der sonomorphologischen Kriterien nach Stavros)	Sensitivität MS für benigne Befunde: 92%			
	Referenzstandard: FNA	Sensitivität MS für maligne Befunde: 67%			
Beurteilung verblindet: -	43 x Fibroadenom 3 x Riesenfibroadenom 6 x fibrozystische Veränderung 1 x Lipom 2 x DCIS 1 x medulläres Karzinom				
Morrow et al., 1998, USA	Erhebung: retrospektive Kohortenstudie an einem Zentrum	FP MS: (mit Biopsie) (n= 40/197) = 20,3%	MS PPV und NPV kann aus den Angaben nicht berechnet werden 12% Familienanamnese 1. Grades 33% Familienanamnese insgesamt +. Selbst getastete Veränderung: 484 (80%), davon klinisch verifiziert: 36%	Empfehlen Bildgebung bei Pat. nur bei unsicherem klinischen Tastbefund wegen hoher falsch positiver Rate bei unauffälligem oder als positiv eingestuftem Tastbefund	3b (-)
	Population: n= 605 Pat < 40 Jahre (med. Alter 29 J., 17-39) untersucht wegen angebl. palpablem Befund	FP MM: (mit Biospie) (n=22/208) = 10,6%			
	Zeitraum: 2/1994-2/1996	MM+/-MS an n=301 Pat. mit unauffälligem oder benigne erscheinendem Tastbefund führt zu n= 44 zusätzlichen Biopsien aufgrund auffälliger Bildgebung, davon +2 DCIS durch Mikrokalk entdeckt, durch MS keine zusätzlichen			
	Befunder: n= 4				
	Erfahrung: alle erfahren in Brustdiagnostik				
	MS-Frequenz: 7,5-13 Mhz				
	Scoring-System angegeben: nein				

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Referenzstandard: histologische Sicherung und klinisches Follow-up med. 18 Monate an 314 Pat. ohne Biopsie (20 lost) Beurteilung verblindet: nein	Karzinome <u>Biopsien an n= 211 klinisch verifizierten palpablen Befunden:</u> 5% Karzinom (n= 28) 10% Zysten 44% Fibroadenome Follow-up erbringt keine zusätzlichen Karzinome	vom Primary Care Provider entdeckte: 108 (20%) davon verifizierte: 29% (nicht stat. signifikant) FNA Befunde nicht berichtet		
Osako et al., 2007, Japan	Erhebung: retrospektive Kohortenstudie an einem Zentrum Population: n= 165 v. 186 Pat. 30-39 J. med. Alter 35,4 J. mit invasivem Krebs und DCIS 90% (148) mit Symptomen Zeitraum: 1/2001-12/2003 Befunder: n= ? Erfahrung: erfahrene Befunder MS-Frequenz: 7,5 –10 Mhz Scoring-System angegeben: Unterteilung in ‚mass-forming-lesions‘ und non-mass-forming-lesions, weiteres Scoring nicht angegeben Referenzstandard: histologische Sicherung Beurteilung verblindet: nein	Sensitivität MS: n= 157/165 95% Sensitivität MM: n= 141/165 85% Invasive Karzinome: Sensitivität MS: n= 144/146 99% Sensitivität MM: n= 124/146 85% DCIS: Sensitivität MS: n= 13/19 68% Sensitivität MM: n=17/19 89%	90% mit ‚dense breast‘ 146 invasive Karzinome (88%) 19 DCIS (12%)	MS wird als Verfahren mit der besseren Detektionsrate präferiert, MM wird als nützlich zur Entdeckung von Mikrokalzifikationen bei DCIS bewertet	
Vargas et al., 2005, USA	Erhebung: retrospektive Kohortenstudie an einem Zentrum Population: n= 296 Pat < 30 Jahre mit MS von n= 542 (med. Alter 24,8; 15-30) untersucht wegen angebl. palpablem Befund Zeitraum: 12/2001-3/2004 Befunder: n= ? Erfahrung: alle erfahren in Brustdiagnostik MS-Frequenz: 7,5-13 Mhz	Sensitivität MS(Befund): 75,9% MS: n= 296 von n=338 palpablen Befunden aufgrund der klinischen Untersuchung n= 224 mit MS-Befund Histo: 72% Fibroadenome 8% fibrozystische Veränderung 6% Phylloides Tumor	Selbst getastete Veränderung: 434 (80%), davon verifiziert durch klinische Untersuchung: 53% vom Primary Care Provider entdeckte: 108 (20%) davon verifizierte:	Standardisierte klinische Untersuchung, MS und Abklärung palpabler oder MS-Befunde durch CNB wird empfohlen	3b (-)

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
	Scoring-System angegeben: nein Referenzstandard: histologische Sicherung durch CNB bei sonographisch sichtbaren oder palpablen Befunden, Verlaufsbeobachtung (med. Zeit nicht angegeben) Beurteilung verblindet: nein	4% Zysten 3% Mastitis 2% Karzinom (4 invasiv-dukta, 1 Phylloides) Standardisierte klinische Untersuchung und MS haben in diesem Kollektiv die gleiche Signifikanz ($p < 0,001$) für das Auffinden eines echten Herdbefunds Im Verlauf keine FN	18% Mammographie nicht eingesetzt		

6. Literatur

6.1 Evidenz Mammasonographie-Screening

6.1.1 Mammasonographie-Screening – eingeschlossene Studien

1. Buchberger W, Niehoff A, Obrist P, Koekkoek-Doll P, Dunser M. Clinically and mammographically occult breast lesions: detection and classification with high-resolution sonography. *Semin Ultrasound CT MR* 2000;21(4):325-36
2. Corsetti V, Ferrari A, Ghirardi M, Bergonzini R, Bellarosa S, Angelini O, Bani C, Ciatto S. Role of ultrasonography in detecting mammographically occult breast carcinoma in women with dense breasts. *Radiol Med (Torino)* 2006;111(3):440-8
3. Crystal P, Strano SD, Shcharynski S, Koretz MJ. Using sonography to screen women with mammographically dense breasts. *AJR* 2003;181(1):177-82
4. Kaplan SS. Clinical utility of bilateral whole-breast US in the evaluation of women with dense breast tissue. *Radiology* 2001;221(3):641-9
5. 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 27,825 patient evaluations. *Radiology* 2002;225(1):165-75
6. Leconte I, Feger C, Galant C, Berliere M, Berg BV, D'Hoore W, Maldague B. Mammography and subsequent whole-breast sonography of nonpalpable breast cancers: the importance of radiologic breast density. *AJR Am J Roentgenol* 2003;180(6):1675-9

6.1.2 Mammasonographie-Screening – ausgeschlossene Einzelstudien

1. Geller BM, Vacek PM, Skelly J, Harvey SC. The use of additional imaging increased specificity and decreased sensitivity in screening mammography. *J Clin Epidemiol* 2005;58(9):942-50
2. O'Driscoll D, Warren R, Mackay J, Britton P, Day NE. Screening with breast ultrasound in a population at moderate risk due to family history. *J Med Screen* 2001;8(2):106-9
3. Ohlinger R, Heyer H, Thomas A, Paepke S, Warm H, Klug U, Frese H, Schulz K, Schimming A, Schwesinger G, Kohler G, Wodny M, Kohlmann T, Grunwald S. Non-palpable breast lesions in asymptomatic women: diagnostic value of initial ultrasonography and comparison with mammography. *Anticancer Res* 2006;26(5B):3943-55

6.1.3 Mammasonographie-Screening – ausgeschlossene Reviews

1. Berg WA. Supplemental screening sonography in dense breasts. *Radiol Clin North Am* 2004;42(5):845-51, vi
2. Delorme S. Mammasonographie und Magnetresonanz-Mammographie als ergänzende Methoden im Mammographiescreening. *Radiologe* 2001;41(4):371-8
3. Elmore JG, Armstrong K, Lehman CD, Fletcher SW. Screening for breast cancer. *JAMA* 2005;293(10):1245-56
4. Gordon PB. Ultrasound for breast cancer screening and staging. *Radiol Clin North Am* 2002;40(3):431-41
5. Mehta TS. Current uses of ultrasound in the evaluation of the breast. *Radiol Clin North Am* 2003;41(4):841-56

6. Smith JA, Andreopoulou E. An overview of the status of imaging screening technology for breast cancer. *Ann Oncol* 2004;15 Suppl 1:118-126
7. Zonderland HM. The role of ultrasound in the diagnosis of breast cancer. *Semin Ultrasound CT MR* 2000;21(4):317-24

6.2 Evidenz Axillasonographie

6.2.1 Axillasonographie: eingeschlossener Review

1. Alvarez S, Anorbe E, Alcorta P, Lopez F, Alonso I, Cortes J. Role of sonography in the diagnosis of axillary lymph node metastases in breast cancer: a systematic review. *AJR Am J Roentgenol* 2006;186(5):1342-8

6.2.2 Axillasonographie: eingeschlossene Einzelstudien

1. Bedrosian I, Bedi D, Kuerer HM, Fornage BD, Harker L, Ross MI, Ames FC, Krishnamurthy S, Edeiken-Monroe BS, Meric F, Feig BW, Akins J, Singletary SE, Mirza NQ, Hunt KK. Impact of clinicopathological factors on sensitivity of axillary ultrasonography in the detection of axillary nodal metastases in patients with breast cancer. *Ann Surg Oncol* 2003;10(9):1025-30
2. Deurloo EE, Tanis PJ, Gilhuijs KG, Muller SH, Kroger R, Peterse JL, Rutgers EJ, Valdes OR, Schultze Kool LJ. Reduction in the number of sentinel lymph node procedures by preoperative ultrasonography of the axilla in breast cancer. *Eur J Cancer* 2003;39(8):1068-73
3. Esen G, Gurses B, Yilmaz MH, Ilvan S, Ulus S, Celik V, Farahmand M, Calay OO. Gray scale and power Doppler US in the preoperative evaluation of axillary metastases in breast cancer patients with no palpable lymph nodes. *Eur Radiol* 2005;15(6):1215-23
4. Mathijssen IM, Strijdhorst H, Kiestra SK, Wereldsma JC. Added value of ultrasound in screening the clinically negative axilla in breast cancer. *J Surg Oncol* 2006;94(5):364-7
5. Nori J, Vanzi E, Bazzocchi M, Bufalini FN, Distanti V, Branconi F, Susini T. Role of axillary ultrasound examination in the selection of breast cancer patients for sentinel node biopsy. *Am J Surg* 2007;193(1):16-20
6. Podkrajsek M, Music MM, Kadivec M, Zgajnar J, Besic N, Pogacnik A, Hocevar M. Role of ultrasound in the preoperative staging of patients with breast cancer. *Eur Radiol* 2005;15(5):1044-50
7. van Rijk MC, Deurloo EE, Nieweg OE, Gilhuijs KG, Peterse JL, Rutgers EJ, Kroger R, Kroon BB. Ultrasonography and fine-needle aspiration cytology can spare breast cancer patients unnecessary sentinel lymph node biopsy. *Ann Surg Oncol* 2006;13(1):31-5
8. Rajesh YS, Ellenbogen S, Banerjee B. Preoperative axillary ultrasound scan: its accuracy in assessing the axillary nodal status in carcinoma breast. *Breast* 2002;11(1):49-52

6.2.3 Axillasonographie: ausgeschlossene Einzelstudien

1. Brancato B, Zappa M, Bricolo D, Catarzi S, Risso G, Bonardi R, Cariaggi P, Bianchin A, Bricolo P, Rosselli Del TM, Cataliotti L, Bianchi S, Ciatto S. Role of ultrasound-guided fine needle cytology of axillary lymph nodes in breast carcinoma staging. *Radiol Med (Torino)* 2004;108(4):345-55
2. Ciatto S, Brancato B, Risso G, Ambrogetti D, Bulgaresi P, Maddau C, Turco P, Houssami N. Accuracy of fine needle aspiration cytology (FNAC) of axillary lymph nodes as a triage test in breast cancer staging. *Breast Cancer Res Treat* 2007;103(1):85-91
3. Couto D, Dias M, Goncalo M, Pinto E, de Oliveira CF. Diagnostic value of ultrasound and color Doppler in identifying axillary lymph node metastases in patients with breast cancer. *Eur J Gynaecol Oncol* 2004;25(5):568-70

4. Damera A, Evans AJ, Cornford EJ, Wilson AR, Burrell HC, James JJ, Pinder SE, Ellis IO, Lee AH, Macmillan RD. Diagnosis of axillary nodal metastases by ultrasound-guided core biopsy in primary operable breast cancer. *Br J Cancer* 2003;89(7):1310-3
5. Davis JT, Brill YM, Simmons S, Sachleben BC, Cibull ML, McGrath P, Wright H, Romond E, Hester M, Moore A, Samayoa LM. Ultrasound-guided fine-needle aspiration of clinically negative lymph nodes versus sentinel node mapping in patients at high risk for axillary metastasis. *Ann Surg Oncol* 2006;13(12):1545-52
6. de Kanter AY, van Geel AN, Paul MA, van Eijck CH, Henzen-Logmans SC, Kruyt RH, Krenning EP, Eggermont AM, Wiggers T. Controlled introduction of the sentinel node biopsy in breast cancer in a multi-centre setting: the role of a coordinator for quality control. *Eur J Surg Oncol* 2000;26(7):652-6
7. Duchesne N, Jaffey J, Florack P, Duchesne S. Redefining ultrasound appearance criteria of positive axillary lymph nodes. *Can Assoc Radiol J* 2005;56(5):289-96
8. Hashimoto H, Suzuki M, Oshida M, Nagashima T, Yagata H, Shishikura T, Imanaka N, Nakajima N. Quantitative ultrasound as a predictor of node metastases and prognosis in patients with breast cancer. *Breast Cancer* 2000;7(3):241-6
9. Kebudi A, Caliskan C, Yetkin G, Celebi S, Isgor A, Mesrur HA, Ezdesir R, Akgun I. The role of pre-operative B mode ultrasound in the evaluation of the axillary lymph node metastases in the initial staging of breast carcinoma. *Acta Chir Belg* 2005;105(5):511-4
10. Khakpour N, Zager JS, Yen T, Stephens T, Kuerer HM, Singletary ES, Ross MI, Hunt KK, Babiera GV. The role of ultrasound in the surgical management of patients diagnosed with ductal carcinoma in situ of the breast. *Breast J* 2006;12(3):212-5
11. Krishnamurthy S, Sneige N, Bedi DG, Edieken BS, Fornage BD, Kuerer HM, Singletary SE, Hunt KK. Role of ultrasound-guided fine-needle aspiration of indeterminate and suspicious axillary lymph nodes in the initial staging of breast carcinoma. *Cancer* 2002;95(5):982-8
12. Kuenen-Boumeester V, Menke-Pluymers M, de Kanter AY, Obdeijn IM, Urich D, Van Der Kwast TH. Ultrasound-guided fine needle aspiration cytology of axillary lymph nodes in breast cancer patients. A preoperative staging procedure. *Eur J Cancer* 2003;39(2):170-4
13. Lemos S, Dias M, Goncalo M, Pinto E, Fernandes G, Oliveira C. Detection of axillary metastases in breast cancer patients using ultrasound and colour Doppler combined with fine needle aspiration cytology. *Eur J Gynaecol Oncol* 2005;26(2):165-6
14. Lumachi F, Tregnaghi A, Ferretti G, Povolato M, Marzola MC, Zucchetta P, Cecchin D, Bui F. Accuracy of ultrasonography and ^{99m}Tc-sestamibi scintimammography for assessing axillary lymph node status in breast cancer patients. A prospective study. *Eur J Surg Oncol* 2006;32(9):933-6
15. Motomura K, Inaji H, Komoike Y, Kasugai T, Nagumo S, Hasegawa Y, Noguchi S, Koyama H. Gamma probe and ultrasonographically-guided fine-needle aspiration biopsy of sentinel lymph nodes in breast cancer patients. *Eur J Surg Oncol* 2001;27(2):141-5
16. Nori J, Bazzocchi M, Boeri C, Vanzi E, Nori BF, Mangialavori G, Distante V, Masi A, Simoncini R, Londero V. Role of axillary lymph node ultrasound and large core biopsy in the preoperative assessment of patients selected for sentinel node biopsy. *Radiol Med (Torino)* 2005;109(4):330-44
17. Ohta M, Tokuda Y, Saitoh Y, Suzuki Y, Okumura A, Kubota M, Makuuchi H, Tajima T, Yasuda S, Shohtsu A. Comparative efficacy of positron emission tomography and ultrasonography in preoperative evaluation of axillary lymph node metastases in breast cancer. *Breast Cancer* 2000;7(1):99-103
18. Oruwari JU, Chung MA, Koelliker S, Steinhoff MM, Cady B. Axillary staging using ultrasound-guided fine needle aspiration biopsy in locally advanced breast cancer. *Am J Surg* 2002;184(4):307-9
19. Popli MB, Sahoo M, Mehrotra N, Choudhury M, Kumar A, Pathania OP, Thomas S. Preoperative ultrasound-guided fine-needle aspiration cytology for axillary staging in breast carcinoma. *Australas Radiol* 2006;50(2):122-6
20. Sapino A, Cassoni P, Zanon E, Fraire F, Croce S, Coluccia C, Donadio M, Bussolati G. Ultrasonographically-guided fine-needle aspiration of axillary lymph nodes: role in breast cancer management. *Br J Cancer* 2003;88(5):702-6

21. Sato K, Tamaki K, Tsuda H, Kosuda S, Kusano S, Hiraide H, Mochizuki H. Utility of axillary ultrasound examination to select breast cancer patients suited for optimal sentinel node biopsy. *Am J Surg* 2004;187(6):679-83
22. Somasundar P, Gass J, Steinhoff M, Koeliker S, Dizon D, Cady B, Taneja C. Role of ultrasound-guided axillary fine-needle aspiration in the management of invasive breast cancer. *Am J Surg* 2006;192(4):458-61
23. Topal U, Punar S, Tasdelen I, Adim SB. Role of ultrasound-guided core needle biopsy of axillary lymph nodes in the initial staging of breast carcinoma. *Eur J Radiol* 2005;56(3):382-5
24. Yang WT, Chang J, Metreweli C. Patients with breast cancer: differences in color Doppler flow and gray-scale US features of benign and malignant axillary lymph nodes. *Radiology* 2000;215(2):568-73
25. Zgajnar J, Besic N, Podkrajsek M, Hertl K, Frkovic-Grazio S, Hocevar M. Minimal risk of macrometastases in the non-sentinel axillary lymph nodes in breast cancer patients with micrometastatic sentinel lymph nodes and preoperatively ultrasonically uninvolved axillary lymph nodes. *Eur J Cancer* 2005;41(2):244-8
26. Zgajnar J, Hocevar M, Podkrajsek M, Hertl K, Frkovic-Grazio S, Vidmar G, Besic N. Patients with preoperatively ultrasonically uninvolved axillary lymph nodes: a distinct subgroup of early breast cancer patients. *Breast Cancer Res Treat* 2006;97(3):293-9

6.2.4 Axillasonographie - Ausgeschlossene Reviews

1. Balu-Maestro C, Chapellier C, Carrier P, Darcourt J, Ettore F, Raoust I. [Role of imaging in the evaluation of axillary nodes and metastatic disease from breast carcinoma]. *J Radiol* 2005;86(11):1649-57
2. Lernevall A. Imaging of axillary lymph nodes. *Acta Oncol* 2000;39(3):277-81

6.3 Evidenz Mammasonographie bei auffälligem Befund

6.3.1 Mammasonographie bei auffälligem Befund – eingeschlossene Einzelstudien

1. Berg WA, Gutierrez L, NessAiver MS, Carter WB, Bhargavan M, Lewis RS, Ioffe OB. Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer. *Radiology* 2004;233(3):830-49
2. Cid JA, Rampaul RS, Ellis IO, Wilson AR, Burrell HC, Evans AJ, Blamey RW, Macmillan RD. Woman feels breast lump--surgeon cannot: the role of ultrasound in arbitration. *Eur J Cancer* 2004;40(14):2053-5
3. Flobbe K, Bosch AM, Kessels AG, Beets GL, Nelemans PJ, von Meyenfeldt MF, van Engelshoven JM. The additional diagnostic value of ultrasonography in the diagnosis of breast cancer. *Arch Intern Med* 2003;163(10):1194-9
4. Georgian-Smith D, Taylor KJ, Madjar H, Goldberg B, Merritt CR, Bokobsa J, Rubin E, Mendelson EB, Fornage BD, Rouse G, Wadden NA, Dewbury KC, Cosgrove DO, Schmidt R. Sonography of palpable breast cancer. *J Clin Ultrasound* 2000;28(5):211-6
5. Kaiser JS, Helvie MA, Blacklaw RL, Roubidoux MA. Palpable breast thickening: role of mammography and US in cancer detection. *Radiology* 2002;223(3):839-44
6. Moss HA, Britton PD, Flower CD, Freeman AH, Lomas DJ, Warren RM. How reliable is modern breast imaging in differentiating benign from malignant breast lesions in the symptomatic population? *Clin Radiol* 1999;54(10):676-82
7. Rahbar G, Sie AC, Hansen GC, Prince JS, Melany ML, Reynolds HE, Jackson VP, Sayre JW, Bassett LW. Benign versus malignant solid breast masses: US differentiation. *Radiology* 1999;213(3):889-94
8. Shetty MK, Shah YP, Sharman RS. Prospective evaluation of the value of combined mammographic and sonographic assessment in patients with palpable abnormalities of the breast. *J Ultrasound Med* 2003;22(3):263-8

9. Skaane P, Sager EM, Olsen JB, Abdelnoor M, Berger A, Wolff PA, Kullmann G. Diagnostic value of ultrasonography in patients with palpable mammographically noncalcified breast tumors. *Acta Radiol* 1999;40(2):163-8
10. Skaane P. The additional value of US to mammography in the diagnosis of breast cancer. A prospective study. *Acta Radiol* 1999;40(5):486-90
11. Zonderland HM, Coerkamp EG, Hermans J, van d, V, van Voorthuisen AE. Diagnosis of breast cancer: contribution of US as an adjunct to mammography. *Radiology* 1999;213(2):413-22

6.3.2 Mammasonographie bei auffälligem Befund – ausgeschlossene Einzelstudien

1. Arger PH, Sehgal CM, Conant EF, Zuckerman J, Rowling SE, Patton JA. Interreader variability and predictive value of US descriptions of solid breast masses: pilot study. *Acad Radiol* 2001;8(4):335-42
2. Baez E, Strathmann K, Vetter M, Madjar H, Hackeloer BJ. Likelihood of malignancy in breast lesions characterised by ultrasound with a combined diagnostic score. *Ultrasound Med Biol* 2005;31(2):179-84
3. Berg WA, Gilbreath PL. Multicentric and multifocal cancer: whole-breast US in preoperative evaluation. *Radiology* 2000;214(1):59-66
4. Beyer T, Moonka R. Normal mammography and ultrasonography in the setting of palpable breast cancer. *Am J Surg* 2003;185(5):416-9
5. Cha JH, Moon WK, Cho N, Chung SY, Park SH, Park JM, Han BK, Choe YH, Cho G, Im JG. Differentiation of benign from malignant solid breast masses: conventional US versus spatial compound imaging. *Radiology* 2005;237(3):841-6
6. Chala L, Endo E, Kim S, de CF, Moraes P, Cerri G, de BN. Gray-scale sonography of solid breast masses: diagnosis of probably benign masses and reduction of the number of biopsies. *J Clin Ultrasound* 2007;35(1):9-19
7. Dennis MA, Parker SH, Klaus AJ, Stavros AT, Kaske TI, Clark SB. Breast biopsy avoidance: the value of normal mammograms and normal sonograms in the setting of a palpable lump. *Radiology* 2001;219(1):186-91
8. Greene T, Cocilovo C, Estabrook A, Chinitz L, Giuliano C, Rosenbaum SS, Tartter PI. A single institution review of new breast malignancies identified solely by sonography. *J Am Coll Surg* 2006;203(6):894-8
9. Hong AS, Rosen EL, Soo MS, Baker JA. BI-RADS for sonography: positive and negative predictive values of sonographic features. *AJR Am J Roentgenol* 2005;184(4):1260-5
10. Mitnick JS, Gianutsos R, Pollack AH, Susman M, Baskin BL, Ko WD, Pressman PI, Feiner HD, Roses DF. Tubular carcinoma of the breast: sensitivity of diagnostic techniques and correlation with histopathology. *AJR* 1999;172(2):319-23
11. Moy L, Slanetz PJ, Moore R, Satija S, Yeh ED, McCarthy KA, Hall D, Staffa M, Rafferty EA, Halpern E, Kopans DB. Specificity of mammography and US in the evaluation of a palpable abnormality: retrospective review. *Radiology* 2002;225(1):176-81
12. Soo MS, Rosen EL, Baker JA, Vo TT, Boyd BA. Negative predictive value of sonography with mammography in patients with palpable breast lesions. *AJR Am J Roentgenol* 2001;177(5):1167-70
13. Taylor KJ, Merritt C, Piccoli C, Schmidt R, Rouse G, Fornage B, Rubin E, Georgian-Smith D, Winsberg F, Goldberg B, Mendelson E. Ultrasound as a complement to mammography and breast examination to characterize breast masses. *Ultrasound Med Biol* 2002;28(1):19-26

6.3.3 Mammasonographie bei auffälligem Befund – ausgeschlossene Reviews

1. Harvey JA. Sonography of palpable breast masses. *Semin Ultrasound CT MR* 2006;27(4):284-97
2. Hille H, Vetter M, Hackeloer BJ. Neubewertung der Mammasonographie in der gegenwärtigen Mammakarzinomdiagnostik. *Ultraschall Med* 2004;25(6):411-7

3. Pruthi S. Detection and evaluation of a palpable breast mass. *Mayo Clin Proc* 2001;76(6):641-7

6.4. Mammasonographie bei Frauen unter 50 Jahren – eingeschlossene Studien

1. Di Nubila B, Cassano E, Urban LA, Fedele P, Abbate F, Maisonneuve P, Veronesi P, Renne G, Bellomi M. Radiological features and pathological-biological correlations in 348 women with breast cancer under 35 years old. *Breast* 2006;15(6):744-53
2. Foxcroft LM, Evans EB, Hirst C, Hicks BJ. Presentation and diagnosis of adolescent breast disease. *Breast* 2001;10(5):399-404
3. Foxcroft LM, Evans EB, Porter AJ. The diagnosis of breast cancer in women younger than 40. *Breast* 2004;13(4):297-306
4. Houssami N, Irwig L, Loy C. Accuracy of combined breast imaging in young women. *Breast* 2002;11(1):36-40
5. Houssami N, Ciatto S, Irwig L, Simpson JM, Macaskill P. The comparative sensitivity of mammography and ultrasound in women with breast symptoms: an age-specific analysis. *Breast* 2002;11(2):125-30
6. Houssami N, Irwig L, Simpson JM, McKessar M, Blome S, Noakes J. Sydney Breast Imaging Accuracy Study: Comparative sensitivity and specificity of mammography and sonography in young women with symptoms. *AJR Am J Roentgenol* 2003;180(4):935-40
7. Malik G, Waqar F, Buledi GQ. Sonomammography for evaluation of solid breast masses in young patients. *J Ayub Med Coll Abbottabad* 2006;18(2):34-7
8. Morrow M, Wong S, Venta L. The evaluation of breast masses in women younger than forty years of age. *Surgery* 1998;124(4):634-40
9. Osako T, Takahashi K, Iwase T, Iijima K, Miyagi Y, Nishimura S, Tada K, Makita M, Akiyama F, Sakamoto G, Kasumi F. Diagnostic ultrasonography and mammography for invasive and noninvasive breast cancer in women aged 30 to 39 years. *Breast Cancer* 2007;14(2):229-33
10. Vargas HI, Vargas MP, Eldrageely K, Gonzalez KD, Burla ML, Venegas R, Khalkhali I. Outcomes of surgical and sonographic assessment of breast masses in women younger than 30. *Am Surg* 2005;71(9):716-9

6.5. Anteil maligner Befunde bei BIRADS 3 – Studien zum Thema

1. Berube M, Curpen B, Ugolini P, Lalonde L, Ouimet-Oliva D. Level of suspicion of a mammographic lesion: use of features defined by BI-RADS lexicon and correlation with large-core breast biopsy. *Can Assoc Radiol J* 1998;49(4):223-8
2. Buchbinder SS, Leichter IS, Lederman RB, Novak B, Bamberger PN, Sklair-Levy M, Yarmish G, Fields SI. Computer-aided classification of BI-RADS category 3 breast lesions. *Radiology* 2004;230(3):820-3
3. Burns RP, Brown JP, Roe SM, Sprouse LR, Yancey AE, Witherspoon LE. Stereotactic core-needle breast biopsy by surgeons: minimum 2-year follow-up of benign lesions. *Ann Surg* 2000;232(4):542-8
4. Caplan LS, Blackman D, Nadel M, Monticciolo DL. Coding mammograms using the classification "probably benign finding - short interval follow-up suggested". *AJR Am J Roentgenol* 1999;172(2):339-42
5. Gokalp G, Topal U. MR imaging in probably benign lesions (BI-RADS category 3) of the breast. *Eur J Radiol* 2006;57(3):436-44
6. Graf O, Helbich TH, Fuchsjäger MH, Hopf G, Morgun M, Graf C, Mallek R. Sonographische Verlaufskontrollen bei palpablen soliden, wahrscheinlich gutartigen Brustläsionen (BI-RADS Kategorie III). *Rofo* 2004;176(9):1251-6
7. Hall FM. Malignancy in BI-RADS category 3 mammographic lesions. *Radiology* 2002;225(3):918-9
8. Kerlikowske K, Smith-Bindman R, Ljung BM, Grady D. Evaluation of abnormal mammography results and palpable breast abnormalities. *Ann Intern Med* 2003;139(4):274-84

9. Liberman L, Abramson AF, Squires FB, Glassman JR, Morris EA, Dershaw DD. The breast imaging reporting and data system: positive predictive value of mammographic features and final assessment categories. *AJR Am J Roentgenol* 1998;171(1):35-40
10. Liberman L, Drotman M, Morris EA, LaTrenta LR, Abramson AF, Zakowski MF, Dershaw DD. Imaging-histologic discordance at percutaneous breast biopsy. *Cancer* 2000;89(12):2538-46
11. Lorenzen J, Wedel AK, Lisboa BW, Loning T, Adam G. Diagnostische Mammographie und Sonographie: Korrelation von diagnostischer BI-RADS-Einstufung mit dem histologischen und klinischen Endbefund. *Rofo* 2005;177(11):1545-51
12. Masroor I. Prediction of benignity or malignancy of a lesion using BI-RADS. *J Coll Physicians Surg Pak* 2005;15(11):686-8
13. Mayi-Tsonga S, Meye JF, Ngou-Mve-Ngou JP, Mendome G, Mounanga M. [Nonpalpable breast lesions: Correlation of the BI-RADS classification and histologic findings.]. *Sante* 2006;16(3):179-83
14. Mendez A, Cabanillas F, Echenique M, Malekshamran K, Perez I, Ramos E. Evaluation of Breast Imaging Reporting and Data System Category 3 mammograms and the use of stereotactic vacuum-assisted breast biopsy in a nonacademic community practice. *Cancer* 2004;100(4):710-4
15. Mendez A, Cabanillas F, Echenique M, Malekshamran K, Perez I, Ramos E. Mammographic features and correlation with biopsy findings using 11-gauge stereotactic vacuum-assisted breast biopsy (SVABB). *Ann Oncol* 2004;15(3):450-4
16. Muller-Schimpfle M, Wersebe A, Xydeas T, Fischmann A, Vogel U, Fersis N, Claussen CD, Siegmann K. Microcalcifications of the breast: how does radiologic classification correlate with histology? *Acta Radiol* 2005;46(8):774-81
17. Obenauer S, Fischer U, Baum F, Dammert S, Fuzesi L, Grabbe E. Stereotaktische Vakuumstanzbiopsie zur Abklärung von gruppiertem Mikrokalk der Kategorie BI-RADS 3. *Rofo* 2001;173(8):696-701
18. Orel SG, Kay N, Reynolds C, Sullivan DC. BI-RADS categorization as a predictor of malignancy. *Radiology* 1999;211(3):845-50
19. Poplack SP, Tosteson AN, Grove MR, Wells WA, Carney PA. Mammography in 53,803 women from the New Hampshire mammography network. *Radiology* 2000;217(3):832-40
20. Rosen EL, Baker JA, Soo MS. Malignant lesions initially subjected to short-term mammographic follow-up. *Radiology* 2002;223(1):221-8
21. Sickles EA. Periodic mammographic follow-up of probably benign lesions: results in 3,184 consecutive cases. *Radiology* 1991;179(2):463-8
22. Sickles EA. Management of probably benign breast lesions. *Radiol Clin North Am* 1995;33(6):1123-30
23. Siegmann KC, Wersebe A, Fischmann A, Fersis N, Vogel U, Claussen CD, Muller-Schimpfle M. Stereotaktische Vakuumbiopsie mammographisch detektierter Läsionen - Erfolg, histologische Sicherheit, Patientenakzeptanz und BI-RADS trade mark -korrelierte Indikationsoptimierung. *Rofo* 2003;175(1):99-104
24. Tan YY, Wee SB, Tan MP, Chong BK. Positive predictive value of BI-RADS categorization in an Asian population. *Asian J Surg* 2004;27(3):186-91
25. Tate PS, Rogers EL, McGee EM, Page GV, Hopkins SF, Shearer RG, Harris JM, Johnson RC, Dubilier LD, Bensema MH, Jansen JF, Clark TD. Stereotactic breast biopsy: a six-year surgical experience. *J Ky Med Assoc* 2001;99(3):98-103
26. Travade A, Isnard A, Bagard C, Bouchet F. [Malignancy in ACR 3 category mammographic lesions, masses and calcifications]. *J Radiol* 2004;85(12 Pt 2):2111-22
27. Varas X, Leborgne JH, Leborgne F, Mezzera J, Jaumandreu S, Leborgne F. Revisiting the mammographic follow-up of BI-RADS category 3 lesions. *AJR Am J Roentgenol* 2002;179(3):691-5
28. Vizcaino I, Gadea L, Andreo L, Salas D, Ruiz-Perales F, Cuevas D, Herranz C, Bueno F. Short-term follow-up results in 795 nonpalpable probably benign lesions detected at screening mammography. *Radiology* 2001;219(2):475-83

29. Wiratkapun C, Wibulpolprasert B, Lertsithichai P. Breast cancer in patients initially assigned as BI-RADS category 3. *J Med Assoc Thai* 2006;89(6):834-9
30. Wiratkapun C, Lertsithichai P, Wibulpholprasert B. Positive predictive value of breast cancer in the lesions categorized as BI-RADS category 5. *J Med Assoc Thai* 2006;89(8):1253-9
31. Zonderland HM, Pope TL, Jr., Nieborg AJ. The positive predictive value of the breast imaging reporting and data system (BI-RADS) as a method of quality assessment in breast imaging in a hospital population. *Eur Radiol* 2004;14(10):1743-50

A Mammographie-Screening
bei Frauen 40-49 Jahre und >70 Jahre
Nutzen/Schaden-Bilanz

B Doppelbefundung von Screening-Mammographien
unter Berücksichtigung des
Entscheidungsverfahrens bei diskrepan-
ten Befunden und des Einsatzes von Computer-
Assistierten-Diagnose (CAD)-Systemen

C Stellenwert der Mammasonographie im Rahmen
der Brustkrebsfrüherkennung

D **Indikationsstellung für die
stereotaktische Vakuumbiopsie im
Vergleich zur stereotaktischen
Stanzbiopsie**

E Nutzen der Exzision des Stichkanals nach minimal-
invasiver Biopsie der Brust zur Vermeidung von
Tumorzellverschleppung

F Her-2/neu-Bestimmung:
Vergleich der Methoden Immunhistochemie,
Fluoreszenz-in-situ-Hybridisierung und
Chromogene-in-situ-Hybridisierung

Inhalt

1. Hintergrund	151
2. Fragestellung	152
3. Methoden	152
3.1 Kriterien für den Einschluss- und Ausschluss von Studien.....	152
3.1.1 Zielgruppe	152
3.1.2 Intervention und Referenzstandard	152
3.1.3 Studientypen	152
3.1.4 Erforderliche Angaben.....	152
3.1.5 Einschluss- und Ausschlusskriterien.....	153
3.2 Literaturrecherche	153
3.3 Zielgrößen und deren Definition	153
3.4 Bewertung der gefundenen Literatur	153
4. Ergebnisse	154
4.1 Ergebnisse der Literaturrecherche	154
4.2 Aufbereitete Evidenz.....	154
4.3 Evidenz von Einzelstudien	156
4.4 Vergleich der Komplikationsraten (gepoolte Daten)	157
4.5 Diskussion	157
5. Evidenztabellen.....	159
5.1 Aufbereitete Evidenz.....	159
5.1.1 Meta-Analyse von Fahrbach et al., 2006	159
5.1.2 Meta-Analyse von Deck, 2006	161
5.1.3 Meta-Analyse von Hoorntje et al., 2003	162
5.1.4 Meta-Analyse von Verkojen et al., 2000	163
5.1.5 Ausgeschlossene aufbereitete Evidenz	163
5.2 Evidenz von Einzelstudien	164
5.3 Ausgeschlossene Einzelstudien	169
6. Literatur.....	170

Tabellenverzeichnis

Tabelle 1	Komplikationsraten von CNB und VAB (gepoolte Daten aus 4 Studien)	157
Tabelle 2	Meta-Analyse von Fahrbach et al., 2006	159
Tabelle 3	Meta-Analyse von Deck, 2006	161
Tabelle 4	Meta-Analyse von Hoorntje et al., 2003	162
Tabelle 5	Meta-Analyse von Verkojen et al., 2000	163
Tabelle 6	Ausgeschlossene aufbereitete Evidenz	163
Tabelle 7	Evidenz von Einzelstudien zum Vergleich stereotaktische Stanzbiopsie – stereotaktische Vakuumbiopsie	164
Tabelle 8	Ausgeschlossene Einzelstudien zum Vergleich stereotaktische Stanzbiopsie- stereotaktische Vakuumbiopsie	169

1. Hintergrund

Suspekte Befunde im Rahmen des organisierten Mammographie-Screenings sind nicht palpabel und werden zum Großteil primär mammographisch diagnostiziert.

Zur histologischen Sicherung suspekter mammographischer Befunde hat die minimal-invasive Biopsie die offene Biopsie fast vollständig ersetzt [1,2]. In Deutschland werden hierfür die Stanzbiopsie und seit 1995 auch die Vakuumbiopsie eingesetzt [3]. Bei dieser Methode der minimal-invasiven Biopsie wird das gewonnene Gewebe nicht ‚ausgestanzt‘, sondern (mit 360° -Rotation) angesaugt und dann abgetrennt. Im Vergleich zur Stanzbiopsie kann so ein größeres Gewebavolumen gewonnen werden [4]. Die Lokalisation mammographisch darstellbarer Befunde erfolgt bei beiden Interventionen mammographisch-stereotaktisch [4]. Die stereotaktische Vakuumbiopsie ist um ein Mehrfaches teurer wie die stereotaktische Stanzbiopsie. Die Materialkosten für die Durchführung einer Biopsie liegen etwa zehn Mal so hoch [5].

Zum Einsatz der stereotaktischen Vakuumbiopsie liegt für Deutschland ein interdisziplinärer Konsensus von 2002 vor [5]. Hierin werden unter Berücksichtigung der Literatur bis 2001 als Indikation mammographische Mikrokalkbefunde und kleine, nicht palpable Herdbefunde genannt. Als eher nicht geeignet für diese Methode werden mammographische Architekturstörungen (v.a. radiäre Narben) und hautnahe Befunde aufgeführt.

In der Leitlinie zur Brustkrebsfrüherkennung von 2003 wird die Entscheidung zur Auswahl der minimal-invasiven Biopsie-Methode weitgehend dem Kliniker überlassen [6].

Durch eine Evidenzrecherche sollte die Frage nach der Indikationsstellung für die stereotaktische Vakuumbiopsie im Vergleich zur stereotaktischen Stanzbiopsie beantwortet werden.

Literatur

1. Kooperationsgemeinschaft Mammographie. Abschlussbericht Modellprojekte (Zusammenfassung). 2006. [cited: 2007 Aug 24]. Available from: http://www.kooperationsgemeinschaft-mammographie.de/cms_upload/datenpool/abschlussberichtmodellprojekte0610.pdf
2. Duijm LE, Groenewoud JH, Roumen RM, de Koning HJ, Plaisier ML, Fracheboud J A. Decade of breast cancer screening in The Netherlands: trends in the preoperative diagnosis of breast cancer. *Breast Cancer Res Treat.* 2007 Jan 12 [im Druck]
3. Lange DG. Vakuumbiopsie der Brust: Erste Erfahrungen (Retrospektive Auswertung der Vakuumbiopsien des Zeitraumes 1996/1997) in der Universitäts-Klinik und Poliklinik für Diagnostische Radiologie an der Martin-Luther-Universität Halle-Wittenberg. Dissertation. Urn:nbn:de:gbv:3-000007248, Universitätsbibliothek Halle, 2004
4. Heywang-Köbrunner SH, Schreer I. Bildgebende Mammadiagnostik. Stuttgart: Thieme Verlag; 2003
5. Obenauer S, Fischer U, Baum F, Grabbe E. Indikationen zur perkutanen stereotaktischen Vakuumbiopsie der Mamma. *Radiologe* 2002;42:11-18.
6. Heywang-Köbrunner SH, Schreer I, Decker T, Böcker W. Interdisziplinärer Konsensus zur Anwendung und Technik der stereotaktischen Vakuumbiopsie. 2002 [cited: 2007 Aug 24] Available from: <http://www.senologie.org/download/pdf/vakuumbiopsie.pdf>
7. Schulz KD, Albert US (Hrsg.). Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland. München: Zuckschwerdt; 2003

2. Fragestellung

Folgende Schlüsselfragen wurden von der Methodikerin in Abstimmung mit dem Auftraggeber formuliert:

1. Für welche mammographischen Befunde ist die stereotaktische Vakuumbiopsie der stereotaktischen Stanzbiopsie hinsichtlich der diagnostischen Sicherheit überlegen?
2. Welche Unterschiede hinsichtlich Art und Umfang von Komplikationen treten bei den beiden Verfahren auf?

3. Methoden

3.1 Kriterien für den Einschluss- und Ausschluss von Studien

Folgende Kriterien wurden von der Methodikerin in Absprache mit der Expertengruppe minimal-invasive Interventionen vorab festgelegt:

3.1.1 Zielgruppe

Zielgruppe waren Frauen mit nicht-palpablen, mammographisch diagnostizierten suspekten Brusttumoren.

3.1.2 Intervention und Referenzstandard

Die Intervention war die stereotaktische Vakuumbiopsie. Der Referenzstandard war die histologische Gewebediagnose oder die mammographische bzw. klinische Kontrolle mit einer Nachbeobachtungszeit von mindestens zwölf Monaten. Die Vergleichsbehandlung war die stereotaktische Stanzbiopsie. Sonographisch gesteuerte Vakuum- oder Stanzbiopsie waren nicht Teil der Intervention oder der Vergleichsbehandlung.

3.1.3 Studientypen

Der methodisch hochwertigste Beleg für die diagnostische Sicherheit der Vakuumbiopsie im Vergleich zur Stanzbiopsie ist durch den Nachweis randomisierter kontrollierter Studien gegeben.

Multizentrische prospektive Kohortenstudien mit konsekutiver Rekrutierung von Patientinnen sind hinsichtlich der Güte der Evidenz ebenfalls als valide einzustufen. Für die vorliegende Fragestellung wurden auch retrospektive Kohortenstudien und Fall-Kontrollstudien berücksichtigt. Systematische Reviews oder Meta-Analysen wurden bei der Bewertung vorrangig berücksichtigt. Recherchiert wurde nach direkten Vergleichsstudien von stereotaktischer Stanzbiopsie zu stereotaktischer Vakuumbiopsie.

3.1.4 Erforderliche Angaben

- Angabe der Gesamtzahl der untersuchten Patientinnen
- Angabe der verwendeten Nadelgröße
- Angabe der Nachbeobachtungszeit
- Angabe des Referenzstandards
(Histologische Sicherung bzw. mammographische/klinische Kontrolle)

3.1.5 Einschluss- und Ausschlusskriterien

Es galten folgende von der Methodikerin festgelegte **Einschlusskriterien:**

1. Studie beantwortet die Fragestellung
2. adäquate Studienpopulation
3. erforderliche Angaben sind genannt

Alle Einschlusskriterien mussten erfüllt sein.

Es galten folgende **Ausschlusskriterien:**

1. Analyse von sonographisch gesteuerter Vakuum- oder Stanzbiopsie
2. Doppelpublikationen

Kein Ausschlusskriterium durfte vorliegen.

3.2 Literaturrecherche

Die systematische Literaturrecherche erfolgte ohne zeitliche Begrenzung. Rechercheschlussdatum war der 28.03.2007.

Nach Vorgabe des Auftraggebers wurden folgende Datenbanken im Rechercheprozess berücksichtigt:

1. PubMed (Internetportal der National Library of Medicine)
(<http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed>)
2. Database of Abstracts of Reviews of Effects (DARE), des Centre for Reviews and Dissemination (CRD)
(<http://www.york.ac.uk/inst/crd/crddatabases.htm#DARE>)
3. Cochrane-Datenbanken 'Cochrane Reviews' und 'Clinical Trials'
(<http://www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/HOME>).

Folgende Suchbegriffe wurden verwendet:

a) Für Pubmed

breast AND (core or high speed or vacuum*) AND (biops*) AND (compar* or complication* or underdiagnosis or underestimat*) AND (stereota* or mammogra*)

b) Für DARE und Cochrane:

breast AND biops* AND vacuum

3.3 Zielgrößen und deren Definition

In Abstimmung mit den Experten wurden folgende Zielgrößen definiert:

- ⇒ Gewinnung von repräsentativem Material
- ⇒ Übereinstimmung von Histologie der Stanze mit der des Tumorexizidats bzw. Korrelation mit Nachbeobachtung (klinisch oder mammographisch mind. 12 Monate)
- ⇒ Komplikationsrate des jeweiligen Verfahrens (differenziert nach Art und Schwere der Komplikation).

3.4 Bewertung der gefundenen Literatur

Für die Evidenzklassifizierung wurde die Oxford-Klassifikation für Diagnose-Studien verwendet (siehe Anhang 1).

4. Ergebnisse

4.1 Ergebnisse der Literaturrecherche

a) In PubMed

Rechercheergebnisse für PubMed	
Anzahl Zitate in Pubmed bis 28.3. 2007	n= 440
Nach Screenen der Abstracts geprüfte Volltexte	n= 13
Reviews	n= 2
Systematische Reviews	n= 2
Einzelstudien, direkte Vergleichsstudien	n= 11
Eingeschlossene Einzelstudien	n= 10

b) In DARE und Cochrane

Rechercheergebnisse für DARE und Cochrane	
Anzahl Zitate (ohne zeitliche Begrenzung bis 28.3.2007)	n= 21
Nach Screenen der Abstracts geprüfte Volltexte	n= 3
Systematische Reviews	n= 1
Health Technology Assessment - Berichte (Meta-Analysen)	n= 2
Eingeschlossene Literatur	n= 1

Durch Handsuche in Literaturverzeichnissen wurden zwei zusätzliche systematische Reviews (Verkooijen et al., 2001; Hetnal et al.) sowie eine Einzelstudie identifiziert (Meyer et al., 1999).

4.2 Aufbereitete Evidenz

a) Studien- und Publikationsqualität

Zum Vergleich der stereotaktischen Stanzbiopsie mit der stereotaktischen Vakuumbiopsie wurden drei Meta-Analysen identifiziert, die die Einschlusskriterien erfüllten (Fahrbach et al., 2006; Deck, 2006 und Hoorntje et al., 2003).

Eine vierte Meta-Analyse – ausschließlich zur diagnostischen Sicherheit der stereotaktischen Stanzbiopsie (Verkooijen et al., 2000) – wird dargestellt, weil sie Teil der Meta-Analyse von Hoorntje et al., 2003 ist. Hoorntje et al. führten Daten zur stereotaktischen Vakuumbiopsie zusammen und verglichen die Ergebnisse dann mit denen von Verkooijen et al.

Die drei Meta-Analysen benutzen jeweils gepoolte Daten von Einzelstudien zur stereotaktischen Vakuumbiopsie und/oder Stanzbiopsie und vergleichen diese. Die gepoolte Studienpopulation aus den Einzelstudien für die jeweilige Intervention liegt zwischen n= 5.004 und n= 6.236.

Für die meisten der Meta-Ergebnisse war der Heterogenitätstest signifikant positiv. Diese Ergebnisse sind im Folgenden deshalb nicht dargestellt. Die drei Meta-Analysen wurden jeweils mit dem Evidenzgrad 3a bewertet.

Vier systematische Reviews bzw. Meta-Analysen wurden ausgeschlossen (zur Begründung siehe siehe 5.1.5 ausgeschlossene aggregierte Evidenz).

In Bezug auf die Publikationsqualität ist anzumerken, dass Fahrbach et al. bei der Ergebnisdarstellung nicht zwischen in situ Karzinomen und invasiven Karzinomen unterscheiden. Hoorntje et al. geben für die Vakuumbiopsie aufgrund ungenügender Nachbeobachtungsdaten in den Studien keinen Anteil nicht erkannter Malignome an. Angaben zur Komplikationsrate der Interventionen fanden sich bei keiner der Meta-Analysen.

b) Inhaltliche Ergebnisse

Testgüte

Der negative prädiktive Wert für die stereotaktische Stanzbiopsie liegt im Median bei 97,4% (97,3% bis 99,2%) und für die stereotaktische Vakuumbiopsie bei 99% und 99,5%.

Die Sensitivität liegt im Median bei 97%, die Spezifität wurde nicht angegeben oder als 100% gesetzt.

Nicht erkannte Karzinome (Biopsieresultat benigne)

Die Rate nicht erkannter Karzinome liegt zwischen 2% und 3,1% für die stereotaktische Stanzbiopsie und zwischen 1,2% und 1,6% für die stereotaktische Vakuumbiopsie.

Anteil an ‚Unterdiagnostik‘

Die drei Meta-Analysen kommen in Bezug auf Biopsiebefunde von atypischer duktaler Hyperplasie zu ähnlichen Ergebnissen.

Die diagnostische Sicherheit für die Erkennung von rein atypischer duktaler Hyperplasie ist für die stereotaktische Vakuumbiopsie deutlich besser als für die stereotaktische Stanzbiopsie. Bei der stereotaktischen Stanzbiopsie ergibt sich ein medianer Anteil von 32,6% (23,3%-38,2%) an in der Biopsie atypischer duktaler Hyperplasie, für die sich am definitiven Operationspräparat das Vorliegen von in-situ-Karzinom oder invasivem Karzinom zeigte und für die stereotaktische Vakuumbiopsie ein medianer Anteil von 17,3% (15,9%-20,4%).

Im Gegensatz dazu sind die Ergebnisse für die Unterdiagnostik nur bezogen auf die invasiven Karzinome im definitiven Operationspräparat, die in der Biopsie als atypische duktale Hyperplasie oder -in-situ-Karzinome befundet wurden, heterogen.

Die gepoolten Daten zeigten bei Deck, 2006 für die stereotaktische Vakuumbiopsie hier 17,7% und für die stereotaktische Stanzbiopsie 11,6%.

Hoorntje et al, 2003 kamen dagegen für die stereotaktische Vakuumbiopsie lediglich auf eine Rate von 11% Unterdiagnostik in Bezug auf die definitive Diagnose eines invasiven Karzinoms bei situ-Karzinom aufgrund der Biopsie. Für die stereotaktische Stanzbiopsie ergaben sich 15% (dieser Unterschied ist statistisch nicht signifikant).

Übereinstimmung Biopsiebefund mit dem definitiven Operationspräparat

Fahrback et al., 2006 ermittelten 97,3% Übereinstimmung für die stereotaktische Vakuumbiopsie und 93,5% für die stereotaktische Stanzbiopsie.

Technisches Versagen

In Bezug auf technisches Versagen bzw. Gewinnung unzureichender Gewebeproben liegen nur von Fahrback et al., 2006 Angaben vor. Die Rate an technischem Versagen liegt für die stereotaktische Vakuumbiopsie bei 1,2% und es wurden keine unzureichenden Gewebeproben gewonnen. Bei der stereotaktischen Stanzbiopsie wurde eine Versagensrate von 5,4% ermittelt und in 2,1% lagen nur unzureichende Gewebeproben vor.

Eine Differenzierung der Ergebnisse nach mammographischem Befund erfolgte in keiner der Analysen, das heißt, es wurde nicht untersucht, ob ein Unterschied in der diagnostischen Sicherheit für bestimmte mammographische Befunde (Mikrokalk, Herdbefunde) zwischen den beiden Verfahren besteht.

4.3 Evidenz von Einzelstudien

a) Studien- und Publikationsqualität

Es wurden zehn retrospektive Kohortenstudien identifiziert, die die Einschlusskriterien erfüllten. Es handelt sich um unizentrische Studien. Die vergleichende Bewertung von stereotaktischer Stanzbiopsie versus stereotaktischer Vakuumbiopsie erfolgte entweder sequentiell, d.h. zu einem bestimmten Zeitpunkt wurde von der einen Intervention auf die andere umgestellt, oder es wurden beide parallel angewendet. Fünf von diesen Studien weisen Gruppengrößen für mindestens eine der beiden Interventionen von unter $n=100$ auf, was die statistische Aussagekraft der Ergebnisse einschränkt.

Die meisten dieser Einzelstudien wurden in den Meta-Analysen – abgesehen vom Publikationsdatum nach Rechercheschluss des jeweiligen Reviews – entweder aufgrund ihrer Größe oder aufgrund der Beschränkung auf die Auswertung spezifischer mammographischer Befunde nicht aufgeführt. Lediglich die Studie von Darling et al., 2000 und Soo et al., 1999 wurde in der Meta-Analyse bei Hoorntje et al. berücksichtigt (ohne befundspezifische Auswertung).

Für die Einzelstudien wurde jeweils ein Evidenzgrad 3b vergeben, da die Befundungen nicht verblindet erfolgten und/oder keine konsekutive Rekrutierung vorlag.

Da sich die Studienpopulation in einigen Studien nur auf positive Befunde bezog, war der negative prädiktive Wert in diesen Studien nicht anzugeben, da der Anteil normal befundeter Biopsien nicht ausgewiesen war.

In Bezug auf die Publikationsqualität ist festzustellen, dass nur in einer Studie über Art und Umfang von Komplikationen berichtet wurde.

b) Inhaltliche Ergebnisse

Für die Analyse von Mikrokalbefunden liegen Daten aus acht Studien vor, für die Analyse von Herdbefunden liegen Daten aus vier Studien vor.

Testgüte bei Mikroverkalkungen

Die Einzelstudien zeigten mehrheitlich eine bessere diagnostische Sicherheit der stereotaktischen Vakuumbiopsie bei dem mammographischen Befund (alleiniger) Mikrokalk:

Meyer et al. berichteten von einem positiven prädiktiven Wert von 91% für die Gewinnung von repräsentativen Mikroverkalkungen bei der stereotaktischen Stanzbiopsie und in 100% bei der stereotaktischen Vakuumbiopsie.

In Bezug auf eine repräsentative Biopsie bei Mikroverkalkungen geben auch Philpotts et al., 1999 für die stereotaktische Vakuumbiopsie einen Zugewinn an positivem prädiktivem Wert von 12% auf 88,3% an.

Ciatto et al. gaben jedoch den negativen prädiktiven Wert in Bezug auf die Erkennung eines invasiven Karzinoms für die stereotaktische Vakuumbiopsie mit 88,9% und für die stereotaktische Stanzbiopsie mit 91,2% an.

In Bezug auf das als falsch-negativ befundene Malignom fanden Ciatto et al. ($n=1.839$) bei isolierten Mikrokalkbefunden 14% mehr falsch-negative Befunde bei der stereotaktischen Stanzbiopsie. In Bezug auf Unterdiagnostik liegt die stereotaktische Vakuumbiopsie im Median insgesamt um 20% (8,7%-38%) niedriger als die stereotaktische Stanzbiopsie (Darling et al.; Wong et al.; Liberman et al.).

Herdbefunde

Vier Studien, bei denen mammographische Herdbefunde untersucht wurden (Soo et al., 2000; Soo et al., 2005; Philpotts et al., 1999; Pfarl et al. 2001) zeigten keinen oder keinen statistisch signifikanten Vorteil der Vakuumbiopsie gegenüber der Stanzbiopsie für diese Befunde. Soo et al. zeigen einen Fall eines nicht erkannten Malignoms und einen Fall von Unterdiagnostik in der stereotaktischen Stanzbiopsie und keine bei stereotaktischer Vakuumbiopsie. Die Gesamtübereinstimmung der Befunde lag für die stereotaktische Stanzbiopsie bei 95% und bei der stereotaktischen Vakuumbiopsie bei 97%. Der negative prädiktive Wert lag bei 98,4% für die stereotaktische Stanzbiopsie und bei 100% für die stereotaktische Vakuumbiopsie.

4.4 Vergleich der Komplikationsraten (gepoolte Daten)

Ein Vergleich der Komplikationsraten war aus den oben aufgeführten Studien nicht möglich, da nur Pfarl et al., 2002 über Komplikationen berichteten.

Für eine repräsentative Beurteilung der Komplikationen wurden deshalb neben Pfarl et al., 2002, bezüglich beider Methoden vier große Einzelstudien verwendet. Insgesamt sind die Angaben zu Komplikationen auch bei großen Studien nicht durchgängig gegeben. Die Art und der Umfang der Komplikationen kann deshalb möglicherweise unterschätzt werden und entsprechend der Vergleich verzerrt sein, da die Genauigkeit der Angaben unterschiedlich sein kann.

Folgende Studien wurden analysiert:

- ⇒ CNB: Fajardo et al., 2004; Meyer et al., 1999; Pfarl et al., 2002
- ⇒ VAB: Kettritz et al., 2004; Beck et al., 2000; Pfarl et al., 2002

Tabelle 1

Komplikationsraten von CNB und VAB (gepoolte Daten aus 4 Studien)

Art der Komplikation	CNB n= 2659 (gepoolt)		VAB n=3846 (gepoolt)	
Hämatom >4cm	Nicht berichtet		n= 25	0,65%
Hämatom mit chirurg. Intervention	0		n= 2	0,05%
Pers. Blutung mit stat. Aufenthalt und / oder chirurgischer Intervention	n= 1	0,03%	n= 7	0,18%
Infektion	n= 1	0,03%	n= 5	0,13%
Pneumothorax	n= 1	0,03%	0	
Relevante Narbenbildung (ggfls Beeinträchtigung der mammographischen Beurteilbarkeit)	-		n= 4	0,1%
Vasovagale Synkope	n= 2	0,07%	n= 6	0,15%
Krampfanfall	0		n= 2	0,05%
Schmerzen oder Unwohlsein, die zum Abbruch der Intervention führten	n= 9	0,3%	0	
Blutung, die zum Abbruch der Intervention führte	n= 3	0,1%	0	
Komplikationsrate gesamt:	0,6%		1,3%	(ohne Hämatome > 4cm): 0,7%

Die Komplikationsraten sind nur eingeschränkt valide vergleichbar, da zum Teil unterschiedliche Kategorien angewendet wurden. Die angegebenen Gesamtkomplikationsraten sind mit 0,6% und 1,3% gering.

Für die Vakuumbiopsie wurden mehr vasovagale Synkopen, persistierende Blutungen mit Komplikationen und Infektionen berichtet sowie zwei Krampfanfälle.

Für die stereotaktische Stanzbiopsie wurde häufiger ein Abbruch der Intervention berichtet.

4.5. Diskussion

a) Methodische Diskussion

Es liegen nach der beschriebenen Recherchestrategie keine Quellen direkter aggregierter Evidenz zu der Frage vor, welche (nicht palpablen) mammographischen Befunde stereotaktisch eher mit stereotaktischer Vakuumbiopsie als mit stereotaktischer Stanzbiopsie abgeklärt werden sollten. Die drei identifizierten Meta-Analysen vergleichen die beiden Verfahren hinsichtlich ihrer diagnostischen Gesamtsicherheit bei nichtpalpablen, mammographisch diagnostizierten Befunden. Es werden jeweils gepoolte Daten von Studien entweder zum einen oder zum anderen Verfahren verglichen. Methodisch sind die drei Meta-Analysen aufgrund der Heterogenität der eingeschlossenen Studien und des indirekten Vergleichs sicher kritisch zu betrachten. In wenigen Einzelstudien werden die beiden Verfahren im Hinblick auf die Ergebnisse unterschiedlicher mammographischer Befunde direkt verglichen. Die Hälfte dieser Studien ist aufgrund einer eher kleinen Fallzahl im Hinblick auf die beobachteten Unterschiede nur eingeschränkt aussagefähig.

Angaben zu den Komplikationen der Verfahren konnten nur einer der vorliegenden Studien entnommen werden, so dass auf eine gepoolte Auswertung großer Einzelstudien zum jeweiligen Verfahren

zurückgegriffen wurde. Die Ergebnisse sind aufgrund unterschiedlicher Komplikationskriterien und aufgrund eventuell lückenhafter Angaben nur als eingeschränkt valide zu bewerten.

b) Inhaltliche Diskussion

Die Frage nach der Indikationsstellung für die stereotaktische Vakuumbiopsie in Bezug auf bestimmte mammographische Befunde kann nur aus Einzelstudien beantwortet werden. Im Vergleich stereotaktische Stanzbiopsie und stereotaktische Vakuumbiopsie ergibt sich für den mammographischen Befund Mikrokalk in der Zusammenschau einen deutlichen Vorteil für die Vakuumbiopsie, nicht aber bei Herdbefunden. Bei Mikrokalk zeigt sich sowohl für die Rate falsch negativer Befunde wie auch für die Rate an Unterdiagnostik eine Differenz von jeweils über 10% zugunsten der stereotaktischen Vakuumbiopsie.

Die vorliegenden Meta-Analysen vergleichen die diagnostische Sicherheit der stereotaktischen Stanzbiopsie und der stereotaktischen Vakuumbiopsie für nichtpalpable mammographische Befunde ohne Analyse der mammographisch vorliegenden Morphologie der Befunde.

Die Rate falsch-negativer Befunde, das bedeutet das Verkennen eines malignen Befundes als gutartig, ist bei beiden Verfahren gering (bis zu 3,1%), die stereotaktische Vakuumbiopsie bietet hier einen Vorteil von 0,8% bis 1,6%.

Im Hinblick auf die Unterdiagnostik von Malignomen (in-situ-Karzinom und invasives Karzinom) aus Befunden atypischer duktaler Hyperplasie zeigen die Meta-Analysen ein übereinstimmendes Ergebnis. Die stereotaktische Vakuumbiopsie bietet gegenüber der stereotaktischen Stanzbiopsie hier einen deutlichen Vorteil von im Median 16,8%. Die klinische Relevanz dieses Befunds kann jedoch als zweifelhaft bewertet werden, da alle Befunde atypischer duktaler Hyperplasie weiterhin zwingend operativ abgeklärt werden sollen. Denn auch mit der Methode stereotaktische Vakuumbiopsie kann bei einem Befund von atypischer duktaler Hyperplasie nicht sicher ausgeschlossen werden, dass ein zusätzliches in-situ-Karzinom vorliegt. In Bezug auf die Unterdiagnostik eines invasiven Karzinoms bei histologischer Diagnose einer atypischer duktaler Hyperplasie oder eines in-situ-Karzinoms aus dem Biopsiegewebe sind die Ergebnisse hinsichtlich der Diagnosesicherheit der beiden Biopsieverfahren widersprüchlich. Eine abschließende Wertung ist deshalb nicht möglich.

Die Komplikationsraten der beiden Verfahren sind nur eingeschränkt valide vergleichbar, da zum Teil unterschiedliche Kategorien angewendet wurden. Die angegebenen Gesamtkomplikationsraten sind mit 0,6% und 1,3% gering.

Für die Vakuumbiopsie wurden mehr vasovagale Synkopen, persistierende Blutungen mit Komplikationen und Infektionen berichtet, sowie zwei Krampfanfälle. Für die stereotaktische Stanzbiopsie wurde häufiger ein Abbruch der Intervention berichtet.

Aufgrund der Datenlage kann die Robustheit der Ergebnisse nicht als valide bezeichnet werden.

Abschließend kann festgestellt für die Gesamtheit mammographischer Befunde festgestellt werden, dass die Ergebnisse, bezogen auf den negativen prädiktiven Wert (jeweils über 97%) und die Gesamtsicherheit der Diagnostik von beiden Interventionen nur wenige Prozentpunkte auseinanderliegen.

5. Evidenztabelle

5.1 Aufbereitete Evidenz

5.1.1 Meta-Analyse von Fahrbach et al., 2006

Tabelle 2

Meta-Analyse von Fahrbach et al., 2006

Fragestellung	Vergleich der diagnostischen Sicherheit und der (technischen) Versagens-Rate von Vakuum-Biopsie im Vergleich zu automatisierter Stanzbiopsie unter stereotaktischer Bildgebung
Datenbanken/ Suchstrategie	Medline, 1996-9/2004 Breast diagnosis/Diagnosis (MeSH) Cytodiagnosis (MeSH) Limits: Human/English und Zeitraum Sowie Suche nach (nicht angegebenen) Keywords als Textwörter um auch nicht verschlagwortete Texte zu finden Zusätzlich Suche im Cochrane Register of controlled Trials
Einschlusskriterien	<ul style="list-style-type: none"> ▪ Englische Sprache ▪ Jegliches Studiendesign mit wenigstens 10 Patienten ▪ Zeitraum (s.o.) ▪ Unselektierte Population (all-comers) in westl. Gesundheitssystemen bei denen nach Screening-Mammographie eine Biopsie indiziert wurde ▪ Überwiegend mammographisch-stereotaktische Bildgebung eingesetzt
Ausgewählte Studien	n= 12 VAB, Anzahl Pat. ges. 5.119 n= 25 CNB, Anzahl Pat. ges. 6.236 Die Autoren identifizierten keine direkte Vergleichsstudie

Ergebnisse	<p>Anteil benigner Befunde, ADH mit und ohne Atypien, sowie invasive Karzinome in beiden Studienpopulationen etwa gleich, bei VAB 7% mehr DCIS. Benigne insgesamt: 67% und 68%. Maligne insgesamt: 33% und 32%</p> <p>Die mammographisch diagnostizierten Läsionen werden nicht differenziert (Herdbefunde, Mikrokalzifikationen).</p> <table border="1" data-bbox="504 448 1390 1211"> <thead> <tr> <th colspan="3">1. Gepoolte Daten zu den 2 Verfahren</th> </tr> <tr> <th></th> <th>VAB</th> <th>CNB</th> </tr> </thead> <tbody> <tr> <td>technisches Versagen:</td> <td>1,5% (58/3916)</td> <td>5,7% (161/2808)</td> </tr> <tr> <td>unzureichende Gewebeproben:</td> <td>0% (0/3665)</td> <td>2,1% (92/4413)</td> </tr> <tr> <td>benigne Läsion mit upgrade maligne:</td> <td>1,2% (29/2452)</td> <td>2% (70/3441)</td> </tr> <tr> <td></td> <td>NPV: 98,8%</td> <td>NPV: 98%</td> </tr> <tr> <td>benigne oder atypische L. mit upgrade maligne:</td> <td>3% (80/2702)</td> <td>5% (186/3745)</td> </tr> <tr> <td>atypische L. mit upgrade maligne:</td> <td>20,4%(51/250)</td> <td>38,2% (116/304)</td> </tr> <tr> <td>Overall Agreement:</td> <td>97,3% (3912/4020)</td> <td>93,5% (5291/5657)</td> </tr> <tr> <th colspan="3">2. Lost of Follow-up</th> </tr> <tr> <th></th> <th>VAB</th> <th>CNB</th> </tr> <tr> <td>Benigne:</td> <td>31,8%</td> <td>20,3%</td> </tr> <tr> <td>mit Atypien:</td> <td>7,8%</td> <td>4,7%</td> </tr> <tr> <td>maligne:</td> <td>2,0%</td> <td>0,7%</td> </tr> <tr> <th colspan="3">3. Meta-Analyse zu Studien mit differenzierter histologischer Aufarbeitung: (Studien nach N gewichtet, Randoms Effect Modell)</th> </tr> <tr> <th></th> <th>VAB</th> <th>CNB</th> </tr> <tr> <td>ADH, upgrade DCIS/IC</td> <td>20,4%</td> <td>38,2%</td> </tr> </tbody> </table>	1. Gepoolte Daten zu den 2 Verfahren				VAB	CNB	technisches Versagen:	1,5% (58/3916)	5,7% (161/2808)	unzureichende Gewebeproben:	0% (0/3665)	2,1% (92/4413)	benigne Läsion mit upgrade maligne:	1,2% (29/2452)	2% (70/3441)		NPV: 98,8%	NPV: 98%	benigne oder atypische L. mit upgrade maligne:	3% (80/2702)	5% (186/3745)	atypische L. mit upgrade maligne:	20,4%(51/250)	38,2% (116/304)	Overall Agreement:	97,3% (3912/4020)	93,5% (5291/5657)	2. Lost of Follow-up				VAB	CNB	Benigne:	31,8%	20,3%	mit Atypien:	7,8%	4,7%	maligne:	2,0%	0,7%	3. Meta-Analyse zu Studien mit differenzierter histologischer Aufarbeitung: (Studien nach N gewichtet, Randoms Effect Modell)				VAB	CNB	ADH, upgrade DCIS/IC	20,4%	38,2%
1. Gepoolte Daten zu den 2 Verfahren																																																				
	VAB	CNB																																																		
technisches Versagen:	1,5% (58/3916)	5,7% (161/2808)																																																		
unzureichende Gewebeproben:	0% (0/3665)	2,1% (92/4413)																																																		
benigne Läsion mit upgrade maligne:	1,2% (29/2452)	2% (70/3441)																																																		
	NPV: 98,8%	NPV: 98%																																																		
benigne oder atypische L. mit upgrade maligne:	3% (80/2702)	5% (186/3745)																																																		
atypische L. mit upgrade maligne:	20,4%(51/250)	38,2% (116/304)																																																		
Overall Agreement:	97,3% (3912/4020)	93,5% (5291/5657)																																																		
2. Lost of Follow-up																																																				
	VAB	CNB																																																		
Benigne:	31,8%	20,3%																																																		
mit Atypien:	7,8%	4,7%																																																		
maligne:	2,0%	0,7%																																																		
3. Meta-Analyse zu Studien mit differenzierter histologischer Aufarbeitung: (Studien nach N gewichtet, Randoms Effect Modell)																																																				
	VAB	CNB																																																		
ADH, upgrade DCIS/IC	20,4%	38,2%																																																		
Methodische Bemerkungen	<p>Die weiteren Ergebnisse der Meta-Analyse wurden hier nicht dargestellt, weil bei allen Studien der Heterogenitätstest hochsignifikant positiv war, d.h. die Inhalte der Studien waren nicht ohne Einschränkung vergleichbar. Insgesamt ist die Rate an Lost of Follow-up bei benignen Läsionen so hoch, dass die Rate nicht erkannter Malignome oder deren Vorstufen einem Bias unterliegen können.</p> <p>Zusatzbemerkung: Die Zahl der im Literaturteil zitierten eingeschlossenen Studien stimmt nicht mit der angegebenen Gesamtzahl überein. Gepoolte Daten sind nicht von allen Studien verfügbar.</p>																																																			
Schlussfolgerung der Autoren	<p>VAB bei im Screening detektierten Läsionen wird CNB gegenüber als das bessere Verfahren gewertet, die Autoren räumen jedoch die methodischen Unsicherheiten ein und empfehlen eine randomisierte Studie zur Bestätigung der Befunde.</p>																																																			
Evidenzgrad	3a																																																			

5.1.2 Meta-Analyse von Deck, 2006

Tabelle 3

Meta-Analyse von Deck, 2006

Fragestellung	Evaluierung der diagnostischen Sicherheit der VAB im Vergleich zu CNB: Soll auf VAB umgestellt werden?
Datenbanken/ Suchstrategie	Medline und INAHTA (zu 3 Zeitpunkten 2005/2006) MeSH: breast, biopsy, vacuum, mammotome, stereota*, percutaneous, needle, large core
Einschlusskriterien	<ul style="list-style-type: none"> ■ 1/1999 –12/2005 ■ n= mind. 100 für jedes Verfahren ■ Vergleich von CNB oder VAB mit Goldstandard FSE oder klinisches langes Follow-up ■ Alle Sprachen
Ausgewählte Studien	n= 8 Studien zu CNB insgesamt 5.616 Pat. (Studie von Fajardo et al bringt 1.174 Pat. ein) n= 12 Studien zu VAB (11g) insgesamt 5.042 Pat. (Studie von Kettritz bringt 2.246 Pat. ein)
Ergebnisse	<p>Miss-Rate Karzinome (Biopsie war benigne, FSE mit Karzinom):</p> <p>CNB 14g: 3,1% NPV: 96,9% VAB 11g: 1,6% NPV: 98,4%</p> <p>Miss-Rate DCIS (Biopsie war ADH, FSE mit DCIS):</p> <p>CNB 14g: 10,7% VAB 11g: 4,0%</p> <p>Miss-Rate IC (Biospie war ADH oder DCIS, FSE mit invasivem Karzinom):</p> <p>CNB 14g: 11,6% VAB 11g: 17,7%</p>
Methodische Bemerkungen	Gepoolte, keine gewichteten Daten! Ergebnis kann einem Bias unterliegen. (in der Studie von Kettritz wird nur 1 invasives Karzinom falsch negativ klassifiziert)
Schlussfolgerung der Autoren	VAB verbessert zwar die Underestimates von DCIS, jedoch nicht die von invasivem Karzinom. VAB soll deshalb CNB nicht grundsätzlich ersetzen, sondern nur für einige klinische Indikationen eingesetzt werden, die erarbeitet werden.
Evidenzbewertung	3a

5.1.3 Meta-Analyse von Hoorntje et al., 2003

Tabelle 4

Meta-Analyse von Hoorntje et al., 2003

Fragestellung	Evaluierung der diagnostischen Sicherheit von VAB und Vergleich mit CNB
Datenbanken/ Suchstrategie	Pubmed 1995-11/2001 breast AND biopsy AND vacuum OR mammotome Zusätzlich Screening der Referenzliteratur
Einschlusskriterien	<ul style="list-style-type: none"> ■ Alle histologischen Biopsiediagnosen sind durch FSE oder adäquates Follow-up gesichert (mind. 90% über 2 Jahre) ■ Absolute Anzahl von benignen und malignen Diagnosen ist ermittelbar ■ Bildgebung ist mammographische Stereotaxie ■ Größe der verwendeten Biopsienadel ist angegeben
Ausgewählte Studien	n=22
Ergebnisse	<p><u>1a. High Risk Underestimate 11g:</u> 15,9% (57/416-57) Fehlendes Follow-up: 13,7% n= 57 Angaben aus 15 Studien; Test für Heterogenität nicht signifikant. Anzahl der Biopsien mit ADH, atypische lobuläre Hyperplasie, LIN, radiäre Narbe, papilläre Läsion oder mögl. Phylloides tumor, die in der FSE in DCIS oder IC upgradet wurden.</p> <p><u>1b. High Risk Underestimate 14g:</u> 23,3% (24/117-14) Fehlendes Follow up: 11,9% n=14 Angaben aus 3 Studien; Test für Heterogenität nicht signifikant</p> <p><u>1c. High Risk Underestimate für Mikrokalzifikationen:</u> 14,5% (12/116-33) Fehlendes Follow up: 28,3% n= 33 Angaben aus 3 Studien: Test für Heterogenität nicht angegeben</p> <p><u>2a. DCIS Underestimate 11g:</u> 10,6% (122/1157-4) Fehlendes Follow up: 0,3% n= 4 Angaben aus 13 Studien; Test für Heterogenität nicht signifikant Anzahl Biopsien mit DCIS, die in der FSE ein Upgrade in IC erfahren.</p> <p><u>2b. DCIS Underestimate 14g:</u> Angaben dazu aus 4 Studien, Heterogenitätstest signifikant, Ergebnis deshalb nicht angegeben</p> <p><u>2c. DCIS Underestimate für Mikrokalzifikationen:</u> 5% und 8% aus 2 Studien mit n= 21 und n= 12</p> <p><u>3. Miss Rates</u> konnten wegen zu großem Anteil fehlenden Follow up bei benignen Biopsien nicht berechnet werden</p> <p><u>4. Vergleich der Underestimates mit einer CNB-Meta-Analyse:</u></p> <p><u>4a. High-Risk:</u> VAB 11g 16% CNB 14g 40% = 24% Diff. (p<0.05)</p> <p><u>4b. DCIS:</u> VAB 11g 11% CNB 14g 15% = 4% Diff. (n. st. sign.)</p> <p><u>Miss Rate CNB:</u> <u>3%</u></p>
Methodische Bemerkungen	Bei gepoolten Daten wurden vor der Auswertung Heterogenitätstests angewandt, diff. der Ergebnisse nach Nadelstärke. Vergleich mit einer CNB Meta-Analyse aus dem Jahr 2000 (Verkojen et al. s.u.)

Schlussfolgerung der Autoren	VAB kann im Vergleich zu CNB High-Risk-Unterdiagnostik verringern, es ist unklar, ob auch die Rate an nicht erkannten Karzinomen verringert werden können. Ergebnisse für Mikrokalzifikationen vergleichbar den 11g Ergebnissen. Ob der Benefit die Kosten aufwiegt, kann aufgrund der fehlenden Raten nicht erkannter Malignome nicht beurteilt werden
Evidenzbewertung	3a

5.1.4 Meta-Analyse von Verkojen et al., 2000

Tabelle 5

Meta-Analyse von Verkojen et al., 2000

Fragestellung	Evaluierung der diagnostischen Sicherheit der 14g CNB bei nicht-palpablen Befunden
Datenbanken/ Suchstrategie	Medline, 1975-Mai 1999 'breast' 'biopsy needle' 'non palpable' 'diagnosis' (Subheadings)
Einschlusskriterien	<ul style="list-style-type: none"> ■ Englische Sprache ■ Non-palpable Läsionen ■ Biopsie gefolgt von FSE oder adäquates Follow up v. 2 Jahren bei 90% der Pat. ■ Entnahme von jeweils mind. 5 Gewebeprobe ■ Anzahl aller benignen und malignen Diagnosen muss angegeben sein
Ausgewählte Studien	n= 5 insgesamt 865 Pat.
Ergebnisse	<p><u>High-Risk-Underestimate:</u> 40% (18/45) Angaben aus 4 Studien mit ins. 763 Pat.; gepoolte Analyse</p> <p><u>DCIS Underestimate:</u> 15% (11/71) Angaben aus 3 Studien mit 571 Pat.; gepoolte Analyse</p> <p><u>Miss Rate:</u> 2,6% (8/307 malign. Diagnosen) NPV: 97,4%</p>
Methodische Bemerkungen	Gepoolte, keine gewichteten Daten. 1 Studie macht etwa 50% der Pat. aus
Schlussfolgerung der Autoren	14g CNB ist aufgrund hoher Sensitivität (97%) und hohem Agreement mit Histo der FSE (94%) eine gute Alternative zur Drahtmarkierung.
Evidenzbewertung	3a

5.1.5 Ausgeschlossene aufbereitete Evidenz

Tabelle 6

Ausgeschlossene aufbereitete Evidenz

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Gutierrez et al., 2005, Spanien	Review zum Einsatz der stereotaktischen Vakuumbiopsie	Volltext nicht frei zugänglich, nur in Spanisch verfügbar
Hetnal et al., 2004, Polen	Meta-Analyse zum Einsatz stereotaktischer Vakuumbiopsie (HTA)	Einschlusskriterien nicht definiert
Corabian, 2005, Canada	Technote zum Einsatz stereotaktischer Vakuumbiopsie (HTA)	Aussagen basieren auf Meta-Analysen von Hoorntje und Hetnal
Medical Services Advisory Committee, 1999, Australien	HTA zum Einsatz der stereotaktischen Vakuumbiopsie im Vergleich zur Core-Needle-Biopsie	Nur Zusammenfassung verfügbar, Literatur nur bis 1998 berücksichtigt

5.2 Evidenz von Einzelstudien

Tabelle 7

Evidenz von Einzelstudien zum Vergleich stereotaktische Stanzbiopsie – stereotaktische Vakuumbiopsie

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Evidenz- grad
Ciatto et al., 2007, Italien	Erhebung:	retrospektiv	Miss-Rate (Karzinom): VAB11g: 23,4% (99/423) CNB14g: 37,4% (47/126) NPV VAB (inv. Ca): 88,9% NPV CNB (inv. Ca): 91,2%	Anzahl fehlender Follow-up's für die jeweilige Methode nicht angegeben Sehr unterschiedl. Gruppengröße	3b
	Population:	n=1223 VAB 11g und 616 CNB 14g aus einer Gesamtzahl von 4035 Biopsien			
	Zeitraum:	1/1996-3/2005			
	Art der Läsion:	isolierte Mikrokalzifikationen			
	Referenzstandard:	histologische Sicherung durch FSE			
Darling et al., USA, 2000	Erhebung:	retrospektiv	Underestimate ADH: CNB 14g: 44% VAB 14g: 39% VAB 11g: 19% Underestimate DCIS: CNB 14g: 21% VAB 14g: 17% VAB 11g: 10% NPV kann nicht berechnet werden	VAB größte Gruppe – beste Expertise	3b
	Population:	n=401 (403 Biopsien) alle mit DCIS oder ADH im Zeitraum von 413 (12 lost of follow up) von insgesamt 3873 B. 67x14gCNB, 75x14gVAB, 261x11gVAB Switch zu VAB 1996			
	Zeitraum:	8/1991-12/1999			
	Art der Läsion:	374x grupp. Mikrokalk 54 Herdbefunde			
	Referenzstandard:	histologische Sicherung durch FSE			
Liberman et al., 2001, USA	Erhebung:	retrospektiv	Underestimate ADH: VAB 11g: 33% (1/3) CNB 14g: 71% (5/7) NPV kann nicht berechnet werden Underestimate DCIS: VAB 11g: 27% (8/30) CNB 14g: 35,7% (5/14) NPV kann nicht berechnet werden	Endgültige Diagnose : DCIS 48,3% (n=43) IC und DCIS 40,4% (n=36) Sehr kleine Gruppengrößen	3b
	Population:	n= 105 von 1341 Läsionen im Zeitraum, 25x14g CNB, 47x11g VAB, 17x14 g VAB, 50 nur FSE			
	Zeitraum:	7/1993-7/2000			
	Art der Läsion:	isolierte, hochsuspekte Mikro-Kalzifikationen BI-RADS 5			
	Referenzstandard:	histologische Sicherung durch FSE			

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Evidenz- grad
Meyer et al., 1997, USA	Erhebung:	retrospektiv	Miss-Rate: (Mikrokalk in den Biopsie-Präparaten nicht enthalten) CNB 14g: 9,2% (n=12) (trotz mehrfacher Entnahme von Zylinder bei fehl. Kalknachweis) <u>Konsequenz:</u> 7 x offene Biopsie (1xDCIS, andere benigne), 4 x Verlaufskontrolle, 1x Re-Biopsie zu späterem Zeitpunkt erfolgreich VAB: 14g: 0% PPV für repräsentative Biopsie: CNB: 91% VAB: 100%	Komplikationsrate als ‚gleich‘ für beide Verfahren beschrieben, n.n. quantifiziert	3b
	Population:	n= 236 alle mit Mikrokalk von 1.032 Biopsien im Zeitraum 130 x 14g CNB und 106 x 14g VAB			
	Zeitraum:	8/1991-12/1996 Switch von CNB zu VAB 1/96			
	Art der Läsion:	gruppiertes Mikrokalk sowohl suspekt als auch benigne imponierend			
	Referenzstandard:	radiologischer Nachweis des Mikrokalks im Biopsiematerial			
Philpotts et al., 1999 USA	Erhebung:	retrospektiv	1. Rebiopsie-Rate <u>a. Aufgrund Atypischer Hyperplasie:</u> 14g CNB 5,2% 11b VAB 5,1% <u>b. Aufgrund insuffizienter Gewebeproben</u> 14g CNB 4,4% 11g VAB 1,7% p=0,042 PPV für suffizientes Material: CNB: 95,6% VAB: 98,6% <u>c. Mammographische-pathologische Differenz:</u> 14g CNB: 3,4% 11g VAB: 0,8%		3b
	Population:	n= 846 (946 Biopsien) n= 521 14g CNB und n= 325 11g VAB			
	Zeitraum:	7/1992-1/1998 (Switch von CNB zu VAB 10/1996)			

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Evidenz- grad
	Art der Läsion:	mammographische Befunde a) CNB-Periode 68% Herdbefunde 32% Mikrokalk b) VAB-Periode 47% Herdbefunde 53% Mikrokalk	p=0.026 <u>d. Rebiopsie für Mikrokalk-Befunde:</u> 14g CNB : 23,7% (45/190) 11g VAB: 11,6% (22/189) p= 0,003 <u>PPV für repräsentative Mikrokalkbiopsie:</u> CNB: 76,3% VAB: 88,3% <u>e. Rebiopsie für Herdbefunde:</u> 14g CNB: 10,7% (43/402) 11g VAB: 6,1% (10/165)		
	Referenzstandard:	histologische Sicherung durch FSE	n.st. sign. <u>PPV für repräsentative Herdbefundbiopsie:</u> CNB: 89,3% VAB: 93,9% <u>f. Gesamt:</u> 14g CNB: 14,9% 11g VAB: 9,0% p=0,009 <u>PPV für repräsentative Biopsie insgesamt:</u> CNB: 85,1% VAB: 91% <u>2. Malignität bei Rebiopsie:</u> 14g CNB: 13,7% 11g VAB: 18,5%		
Soo et al., 2000, USA	Erhebung:	retrospektiv	<u>Underestimate ges:</u> 14g CNB n=1 14g VAB n=0 NPV CNB: 96%	Sehr kleines n!	3b
	Population:	n= 110 (116 Biopsien) n= 73 mit 14g Multipass n= 37 mit 14g VAB			
	Zeitraum:	10/1995 – 8/1997			

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Evidenz- grad
	Art der Läsion:	nicht-palpable, mammographisch entdeckte Befunde ohne Mikrokalk – Herdbefunde, Asymmetrien, Architekturstörungen	NPV VAB: 100% lost of Follow up: 22/110 (20%) Rebiopsie-Rate marginal geringer bei VAB		
Soo et al., 2005, USA	Erhebung:	retrospektiv	<u>Adäquate, mit der FSE konkordante Proben:</u> 14g CNB 95% 14g VAB 97% 11g VAB 97%		3b
	Population:	n= 706 Biopsien gesamt davon n=193 stereotaktische Biopsien n= 76 14g CNB n= 78 14g VAB n= 39 11g VAB	<u>Miss Rate (Karzinom):</u> 14g CNB n=1 14g VAB und 11g VAB n=0. NPV CNB: 98,4% NPV VAB: 100% <u>Kosten-Nutzen-Decision-Tree:</u> <u>Favours 14g CNB</u>		
	Zeitraum:	3 Jahre			
	Art der Läsion:	nicht-palpable, mammographisch entdeckte Befunde ohne Mikrokalk – Herdbefunde, Asymmetrien, Architekturstörungen			
	Referenzstandard:	histologische Sicherung durch FSE oder mammographische Kontrolle			
Won et al., 1999, USA	Erhebung:	retrospektiv	<u>Underestimate DCIS:</u> <u>14g CNB:</u> 35% (n=7) 5x invasives Ca, 2x mikroinvasiv	Sehr kleine Fallzahl	3b
	Population:	n= 40 konsekutive Biopsien mit DCIS von insgesamt 390 im Zeitraum 20x14g CNB, 20x11g VAB			
	Zeitraum:	8/1992-7/1998			
	Art der Läsion:	36 Mikrokalzifikationen, 3 Architekturstörungen, 1 Herdbefund	<u>11g VAB:</u> 15% (n=3) 3x mikroinvasiv		
	Referenzstandard:	histologische Sicherung durch FSE	Unterschied ist nicht statistisch signifikant, p=0.13 NPV kann nicht berechnet werden		

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Evidenz- grad																
Zhao et al., 2003, USA	Erhebung: Population:	retrospektiv n= 1.036 686x VAB 11g 350x CNB 14g davon alle mit ADH n= 53 (15x CNB, 38x VAB)	<u>Underestimate CNB 14g:</u> 4/10 mit Biopsie = 40% (1x IC, 3 DCIS) 67% mit Referenzstandard Biopsie <u>Underestimate VAB 11g:</u> 7/29 mit Biopsie = 24% (nur DCIS) 76% Referenzstandard Biopsie Diff. nicht statistisch signifikant NPV kann nicht berechnet werden	Sehr kleines n von n= 53 ADH waren mammographisch 77% (n=41) Mikrover- kalkungen	3b																
Pfarl et al., 2002, Österreich	Erhebung: Population:	retrospektiv n= 580 an 569 Pat., 168x14g CNB (29%), 134x11g VAB (23%), 278x14g VAB (48%)	52,2% definitiv maligne Läsionen <u>Miss Rate (Karzinom):</u> <table border="1" data-bbox="952 821 1276 941"> <thead> <tr> <th></th> <th>14gC</th> <th>11gV</th> <th>14gV</th> </tr> </thead> <tbody> <tr> <td>a)</td> <td>1,1%</td> <td>2%</td> <td>2,9%</td> </tr> <tr> <td>b)</td> <td>2,8%</td> <td>2,3</td> <td>2%</td> </tr> <tr> <td>c)</td> <td>4,8%</td> <td>3,4</td> <td>0%</td> </tr> </tbody> </table> NPV für Mikrokalzifikationen: NPV für Herdbefunde:		14gC	11gV	14gV	a)	1,1%	2%	2,9%	b)	2,8%	2,3	2%	c)	4,8%	3,4	0%	VAB am besten bei asymmetrischen Verdichtungen und Mikrokalk Komplikationen: 1,2% (n= 7) CNB 14g 1,7% VAB 11g 2,2% (kein signifikanter Unterschied) (s.u.)	3b
	14gC	11gV	14gV																		
a)	1,1%	2%	2,9%																		
b)	2,8%	2,3	2%																		
c)	4,8%	3,4	0%																		
	Zeitraum:	3/1996-4/2002 Switch von CNB zu VAB 1998																			
	Art der Läsion:	nicht-palpable, mammographisch entdeckte Befunde Mikrokalk, Herdbefunde, Asymmetrien, Architekturstörungen																			
	Referenzstandard:	histologische Sicherung durch FSE oder mammographische Kontrolle																			
	Zeitraum:	1994-1999																			
	Art der Läsion:	a) Herdbefunde: 54% (n= 311) b) Mikrokalk: 22% (n= 130) c) Asymm. Verdichtung: 1% (n= 62) d) andere (3 Komb.): 13,3% (n= 77)																			
	Referenzstandard:	histologische Sicherung durch FSE 93,8% oder mammographische Verlaufskontrolle (6,2%)																			

5.3 Ausgeschlossene Einzelstudien

Tabelle 8

Ausgeschlossene Einzelstudien zum Vergleich stereotaktische Stanzbiopsie-
stereotaktische Vakuumbiopsie

Autor, Jahr, Land	Grund für Ausschluss
Sedivy, 1998, Wien	Teilmenge der Studie von Pfarl et al
Velanovich et al., 1999, USA	absolute Zahlenangaben zu den einzelnen Verfahren fehlen

6. Literatur

6.1 Eingeschlossene Reviews/Meta-Analysen

1. Deck W. Macrobiopsie mammaire par aspiration. Montreal: AETMIS; 2006.
2. Fahrbach K, Sledge I, Cella C, Linz H, Ross SD. A comparison of the accuracy of two minimally invasive breast biopsy methods: a systematic literature review and meta-analysis. *Arch Gynecol Obstet* 2006;274(2):63-73.
3. Hoorntje LE, Peeters PH, Mali WP, Borel R, I. Vacuum-assisted breast biopsy: a critical review. *Eur J Cancer* 2003;39(12):1676-83.
4. Verkooijen HM, Peeters PH, Buskens E, Koot VC, Borel R, I, Mali WP, van Vroonhoven TJ. Diagnostic accuracy of large-core needle biopsy for nonpalpable breast disease: a meta-analysis. *Br J Cancer* 2000;82(5):1017-21.

6.2 Ausgeschlossene Reviews/Meta-Analysen

1. Corabian P, Alberta Heritage Foundation for Medical Research (AHFMR). TechNote: Image-Guided Vacuum-Assisted Breast Biopsy for Suspicious, Non-Palpable Breast Lesions. 2005 [cited: 2007 Jun 15]. Available from: http://www.ihe.ca/documents/hta/TechNote__50.pdf
2. Gutierrez A, Taboada J, Apestegua L, Arrizabalaga R, Asua J, Andreu F J Vitoria-Gasteiz, Spain: Basque Office for Health Technology Assessment, Health Department Basque Government (OSTEBA). New percutaneous techniques of histological diagnosis of non palpable lesions suspected of breast cancer.2005
3. Hetnal M, Wechowski J, Wladysiuk-Blicharz M, Landa K. The comparison of efficacy and safety of selected diagnostic methods of non-palpable breast lesions. Cracovie: HTA Consulting; 2005 [cited: 2007 Jun 15]. Available from: <http://www.hta.pl/php/produkty.php?podadres=nowe>
4. Medical Services Advisory Committee, Canberra: Medical Services Advisory Committee (MSAC). Directional, vacuum-assisted breast biopsy. MSAC application 1015: Final Assessment Report. 1999

6.3 Eingeschlossene Einzelstudien CNB-VAB (Vergleichsstudien)

1. Ciatto S, Houssami N, Ambrogetti D, Bianchi S, Bonardi R, Brancato B, Catarzi S, Risso GG. Accuracy and underestimation of malignancy of breast core needle biopsy: the florence experience of over 4000 consecutive biopsies. *Breast Cancer Res Treat* 2007;101(3):291-7
2. Darling ML, Smith DN, Lester SC, Kaelin C, Selland DL, Denison CM, DiPiro PJ, Rose DI, Rhei E, Meyer JE. Atypical ductal hyperplasia and ductal carcinoma in situ as revealed by large-core needle breast biopsy: results of surgical excision. *AJR Am J Roentgenol* 2000;175(5):1341-6
3. Liberman L, Gougoutas CA, Zakowski MF, LaTrenta LR, Abramson AF, Morris EA, Dershaw DD. Calcifications highly suggestive of malignancy: comparison of breast biopsy methods. *AJR Am J Roentgenol* 2001;177(1):165-72
4. Meyer JE, Smith DN, DiPiro PJ, Denison CM, Frenna TH, Harvey SC, Ko WD. Stereotactic breast biopsy of clustered microcalcifications with a directional, vacuum-assisted device. *Radiology* 1997;204(2):575-6

5. Pfarl G, Helbich TH, Riedl CC, Rudas M, Wagner T, Memarsadeghi M, Lomoschitz F. Stereotaktische Nadelbiopsie der Brust: Diagnosesicherheit verschiedener Biopsie-Systeme und Nadelkaliber. *Rofo* 2002;174(5):614-9
6. Philpotts LE, Shaheen NA, Carter D, Lange RC, Lee CH. Comparison of rebiopsy rates after stereotactic core needle biopsy of the breast with 11-gauge vacuum suction probe versus 14-gauge needle and automatic gun. *AJR Am J Roentgenol* 1999;172(3):683-7
7. Soo MS, Ghate S, Delong D. Stereotactic biopsy of noncalcified breast lesions: utility of vacuum-assisted technique compared to multipass automated gun technique. *Clin Imaging* 1999;23(6):347-52
8. Soo MS, Kliewer MA, Ghate S, Helsper RS, Rosen EL. Stereotactic breast biopsy of noncalcified lesions: a cost-minimization analysis comparing 14-gauge multipass automated core biopsy to 14- and 11-gauge vacuum-assisted biopsy. *Clin Imaging* 2005;29(1):26-33
9. Won B, Reynolds HE, Lazaridis CL, Jackson VP. Stereotactic biopsy of ductal carcinoma in situ of the breast using an 11-gauge vacuum-assisted device: persistent underestimation of disease. *AJR Am J Roentgenol* 1999;173(1):227-9
10. Zhao L, Freimanis R, Bergman S, Shen P, Perrier ND, Lesko N, Pulaski T, Pulaski S, Carr JJ, Levine EA. Biopsy needle technique and the accuracy of diagnosis of atypical ductal hyperplasia for mammographic abnormalities. *Am Surg* 2003;69(9):757-62

6.4 Ausgeschlossene Einzelstudien CNB-VAB (Vergleichsstudien)

1. Sedivy R, Partik B, Helbich T, Breitenecker G, Wolf G. Stereotaktische Mammabiopsie: Vergleich von Vakuumstanzbiopsie versus Hochgeschwindigkeits-Corebiopsie. *Wien Med Wochenschr* 1998;148(14):331-4
2. Velanovich V, Lewis FR, Jr., Nathanson SD, Strand VF, Talpos GB, Bhandarkar S, Elkus R, Szymanski W, Ferrara JJ. Comparison of mammographically guided breast biopsy techniques. *Ann Surg* 1999;229(5):625-30

6.5 Studien zur Angabe der Komplikationsraten von CNB und VAB

1. Beck RM, Gotz L, Heywang-Kobrunner SH. Stereotaxic vacuum core breast biopsy-experience of 560 patients. *Swiss Surg* 2000;6(3):108-10
2. Fajardo LL, Pisano ED, Caudry DJ, Gatsonis CA, Berg WA, Connolly J, Schnitt S, Page DL, McNeil BJ. Stereotactic and sonographic large-core biopsy of nonpalpable breast lesions: results of the Radiologic Diagnostic Oncology Group V study. *Acad Radiol* 2004;11(3):293-308
3. Kettritz U, Rotter K, Schreer I, Murauer M, Schulz-Wendtland R, Peter D, Heywang-Kobrunner SH. Stereotactic vacuum-assisted breast biopsy in 2874 patients: a multicenter study. *Cancer* 2004;100(2):245-51
4. Meyer JE, Smith DN, DiPiro PJ, Denison CM, Frenna TH, Harvey SC, Ko WD. Stereotactic breast biopsy of clustered microcalcifications with a directional, vacuum-assisted device. *Radiology* 1997;204(2):575-6
5. Pfarl G, Helbich TH, Riedl CC, Rudas M, Wagner T, Memarsadeghi M, Lomoschitz F. Stereotaktische Nadelbiopsie der Brust: Diagnosesicherheit verschiedener Biopsie-Systeme und Nadelkaliber. *Rofo* 2002;174(5):614-9

A Mammographie-Screening
bei Frauen 40-49 Jahre und >70 Jahre
Nutzen/Schaden-Bilanz

B Doppelbefundung von Screening-Mammographien
unter Berücksichtigung des
Entscheidungsverfahrens bei diskrepanten
Befunden und des Einsatzes von Computer-
Assistierten-Diagnose (CAD)-Systemen

C Stellenwert der Mammasonographie im Rahmen
der Brustkrebsfrüherkennung

D Indikationsstellung für die stereotaktische
Vakuumbiopsie im Vergleich zur stereotaktischen
Stanzbiopsie

E **Nutzen der Exzision des Stichkanals
nach minimal-invasiver Biopsie der
Brust zur Vermeidung von
Tumorzellverschleppung**

F Her-2/neu-Bestimmung:
Vergleich der Methoden Immunhistochemie,
Fluoreszenz-in-situ-Hybridisierung und
Chromogene-in-situ-Hybridisierung

Inhalt

1. Hintergrund	175
2. Fragestellung	176
3. Methoden	176
3.1 Kriterien für den Einschluss und Ausschluss von Studien.....	176
3.1.1 Zielgruppe	176
3.1.2 Intervention und Vergleich.....	176
3.1.3 Studientypen	176
3.1.4 Erforderliche Angaben	176
3.1.5 Ein- und Ausschlusskriterien	177
3.2 Literaturrecherche.....	177
3.3 Zielgrößen und deren Definition.....	177
3.4 Bewertung der gefundenen Literatur.....	178
4. Ergebnisse	178
4.1 Ergebnisse der Literaturrecherche.....	178
4.2 Aufbereitete Evidenz	178
4.3 Evidenz von Einzelstudien	178
4.4 Diskussion.....	179
4.5 Studien zu Tumorzellverschleppung nach minimal invasiver Biopsie	179
4.6 Studien zum Auftreten von Lokalrezidiv nach minimal-invasiver Biopsie	180
5. Evidenztabelle	181
5.1 Evidenz von Einzelstudien	181
5.1.1 Eingeschlossene Einzelstudien.....	181
5.2 Ausgeschlossene Studien	182
5.2.1 Ausgeschlossene aggregierte Evidenz	182
5.2.2 Studien zum Nachweis von Tumorzellverschleppung.....	182
5.2.3 Studien mit dem Endpunkt Rezidiv nach minimal-invasiver Biopsie	184
5.2.4 Fallberichte mit dem Endpunkt Rezidiv nach minimal-invasiver Biopsie	189
6. Literatur.....	190

Tabellenverzeichnis

Tabelle 1	Evidenz von Einzelstudien zum Nutzen der Stichkanalexzision nach minimal-invasiver Biopsie	181
Tabelle 2	Ausgeschlossene Reviews zum Nachweis von Tumorzellverschleppung nach minimal- invasiver Biopsie	182
Tabelle 3	Studien mit Nachweis von Tumorzellverschleppung im exziierten Stichkanal	182
Tabelle 4	Fallberichte zu Tumorzellverschleppung nach minimal-invasiver Biopsie	183
Tabelle 5	Studien mit dem Endpunkt Rezidiv nach minimal-invasiver Biopsie	184
Tabelle 6	Fallberichte zum Auftreten von Rezidiv nach minimal-invasiver Biopsie	189
Tabelle 7	Nicht berücksichtigte Studien zum Rezidiv nach Feinnadelpunktion	189

1. Hintergrund

Zur Diagnosesicherung eines suspekten Brusttumors ist die histologische Abklärung zwingend erforderlich [1;2]. Auch bei nicht-palpablen Brusttumoren wird die Durchführung einer minimal-invasiven Biopsie zur präoperativen histologischen Sicherung empfohlen [1]. Ziel dabei ist nicht nur die Vermeidung offener Biopsien bei möglichen falsch-positiven Befunden, sondern auch die Möglichkeit einer definitiven Operationsplanung und damit die verbesserte präoperative Aufklärung der Patientin [1].

In der Stufe-3-Leitlinie Brustkrebsfrüherkennung in Deutschland, 2003 wird in der Schlüsselempfehlung für die Durchführung perkutaner Biopsien gefordert, dass der Stichkanal bei typischer Schnittführung durch die nachfolgende Operation entfernt werden kann [2]. Diese Forderung gründet sich auf Studien, die nach Brustoperationen am histologischen Präparat Tumorzellverschleppungen nach vorausgegangener minimal-invasiver Biopsie nachwiesen (siehe 5.2.2). Durch die Stichkanalexzision soll eine mögliche Metastasierung aufgrund dieser Tumorzellverschleppung verhindert werden.

Die Datenlage zu der klinischen Relevanz der Tumorzellverschleppung und der Auswirkungen der Stichkanalexzision sollte nun einer aktuellen Evidenzbewertung unterzogen werden.

Literatur

1. Schulz KD, Albert US (Hrsg.). Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland. München: Zuckschwerdt; 2003
2. Deutsche Krebsgesellschaft (Hrsg.). Diagnostik, Therapie und Nachsorge des Mammakarzinoms der Frau. München: Zuckschwerdt; 2004

2. Fragestellung

Folgende Frage wurde von der Methodikerin in Absprache mit der Expertengruppe für minimal-invasive Interventionen formuliert:

1. Ist die Exzision des Stichkanals nach minimal-invasiver Biopsie aufgrund einer möglichen metastastischen Tumorabsiedelung erforderlich?

3. Methoden

3.1 Kriterien für den Einschluss und Ausschluss von Studien

3.1.1 Zielgruppe

Zielgruppe waren Patientinnen mit minimal-invasiver Biopsie. Patientinnen mit Drahtmarkierung bei offener Biopsie waren nicht Teil der Zielgruppe.

3.1.2 Intervention und Vergleich

Die Intervention war die Exzision des Stichkanals nach minimal-invasiver Biopsie und die histologische Untersuchung auf verschleppte Tumorzellen. Die Vergleichsbehandlung war die Belassung des Stichkanals in-situ.

3.1.3 Studientypen

Zur Beantwortung der vorliegenden Fragestellung liefern randomisierte kontrollierte Studien mit dem Endpunkt Lokalrezidiv die besten Argumente um das Fortführen oder Aufgeben der Maßnahme „Exzision des Stichkanals“ zu begründen. Die nächste Stufe in der Evidenzhierarchie nach Oxford bilden für therapeutische Fragestellungen gut geplante Kohortenstudien, danach folgen Outcome-Studien. Darunter wurden Studien verstanden, die zu verschiedenen Zeitpunkten den Outcome nach einer Intervention untersuchen.

Auch Fall-Kontrollstudien und Fallserien wurden in die Bewertung einbezogen. Einzelfallberichte wurden ausgeschlossen (siehe Anhang 1: Oxford-Klassifikation für therapeutische Studien).

3.1.4 Erforderliche Angaben

- Angabe, ob der Stichkanal exzidiert wurde
- Angabe des Operationsverfahrens: Brusterhaltende Therapie oder Mastektomie
- Angabe der postoperativen Bestrahlung
- Med. Nachbeobachtungszeit von mindestens 2 Jahren

3.1.5 Ein- und Ausschlusskriterien

Es galten folgende von der Methodikerin festgelegte **Einschlusskriterien**:

1. Studie beantwortet die Fragestellung
2. adäquate Studienpopulation
3. erforderliche Angaben sind genannt

Alle Einschlusskriterien mussten erfüllt sein

Es galten folgende **Ausschlusskriterien**:

1. ausschließlich Analyse von Feinnadelpunktionen (da in Deutschland kein übliches Verfahren)
2. Doppelpublikationen

Kein Ausschlusskriterium durfte vorliegen.

3.2 Literaturrecherche

Die systematische Literaturrecherche erfolgte ohne zeitliche Begrenzung. Rechercheschlussdatum war der 16.3.2007.

Die Abstracts der jeweiligen Literatursuche wurden von der Methodikerin auf inhaltliche Relevanz geprüft, nicht relevante Abstracts wurden ausgeschlossen. Inhaltlich als relevant beurteilte Abstracts wurden als Volltexte geprüft. Der Ausschluss von Volltexten wird jeweils begründet.

Nach Vorgabe des Auftraggebers erfolgte die Literaturrecherche im ÄZQ in folgenden Datenbanken:

1. PubMed (Internetportal der National Library of Medicine) (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed>)
2. Database of Abstracts of Reviews of Effects (DARE), des Centre for Reviews and Dissemination (CRD) (<http://www.york.ac.uk/inst/crd/crddatabases.htm#DARE>)
3. Cochrane-Datenbanken 'Cochrane Reviews' und 'Clinical Trials' (<http://www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/HOME>).

Aufgrund der geringen Anzahl und Heterogenität der Studien wurde für dieses Thema eine zusätzliche Recherche in den Datenbanken: ‚Embase‘ (Exerpta Medica Database; <http://www.embase.com/>) und Biosis (Biological Science) Previews bis 2004 (<http://scientific.thomson.com/products/bp/>) durchgeführt.

Folgende Suchbegriffe wurden (als Textworte) verwendet:

‚biops* AND (displacement OR seeding) AND breast‘

3.3 Zielgrößen und deren Definition

Für die vorliegende Fragestellung sollte der Nachweis erbracht werden, ob durch Tumorzellverschleppung aufgrund von minimal invasiver Biopsie eine klinisch relevante Metastasierung erfolgt und ob diese durch die Exzision des Stichkanals verhindert werden kann.

Als Zielgrößen wurde deshalb festgelegt:

- ⇒ histologischer Nachweis verschleppter Tumorzellen im Stichkanal
- ⇒ Auftreten eines histologisch gesicherten Lokalrezidivs

Das Lokalrezidiv wurde wie folgt definiert: histologisch gesichertes Rezidiv in der gleichen Brust/auf der gleichen Thoraxseite, bei invasiven Tumoren mit der gleichen Histologie, bei vorausgegangenem in-situ Karzinom auch mit nachgewiesener Invasivität.

3.4 Bewertung der gefundenen Literatur

Die Exzision des Stichkanals ist eine potentiell therapeutische Maßnahme. Für die Evidenzklassifizierung wurde deshalb die Oxford-Klassifikation für therapeutische Studien verwendet (siehe Anhang 1).

4. Ergebnisse

4.1 Ergebnisse der Literaturrecherche

Die Suche in DARE und Cochrane ergab keine Treffer.

Rechercheergebnisse für PubMed	
PubMed bis 16.3.2007	n= 89
Nach Screenen der Abstracts geprüfte Volltexte	n= 29
Reviews	n= 6
Systematische Reviews	n= 0
Studien zur Tumorzellverschleppung am histologischen Präparat	n= 9
Studien zum Auftreten von Rezidiven nach minimal-invasiver Biopsie	n= 11
Handsuche in Lit.verzeichnissen	n= 3 (alle ausgeschl. da zu FNAC)
Eingeschlossene Studien	n= 1

Die zusätzliche Suche in Embase und Biosis ergab keine weiteren Treffer.

4.2 Aufbereitete Evidenz

Es wurde kein systematischer Review zu der fokussierten Fragestellung identifiziert. Sechs Reviews wurden ausgeschlossen (siehe 5.2 Ausgeschlossene Reviews). Fünf Reviews waren narrativ. Ein Review von Smith et al., 1984 diskutierte Ergebnisse zu Tumorzellverschleppung von Feinnadelpunktionen.

4.3 Evidenz von Einzelstudien

a) Studien- und Publikationsqualität

Zu der vorliegenden Fragestellung wurde eine Studie identifiziert, die die Einschlusskriterien erfüllte (Stolier et al., 2000). Stolier et al. exzidierten als einzige Arbeitsgruppe zunächst den Stichkanal und führten einen histologischen Nachweis über verschleppte Tumorzellen. Anschließend wurden die Patientinnen über mehr als zwei Jahre nachbeobachtet.

Alle anderen Einzelstudien erfüllten nicht die Einschlusskriterien zur Beantwortung der fokussierten Frage.

Im Folgenden ist der Ausschluss der weiteren Studien zum Thema Tumorzellverschleppung und Lokalrezidiv nach minimal-invasiver Biopsie begründet.

1. Zum histologischen Nachweis von Tumorzellverschleppungen im Operationspräparat konnten insgesamt sechs weitere Studien (n= 28 bis n= 352) und drei Einzelfallberichte identifiziert werden. Alle Studien untersuchten Fallserien von Operationspräparaten mit exzidiertem Stichkanal. Zwei Studien waren jedoch nicht auf minimal-invasive Biopsien bezogen. Keine dieser Studien nahm zum Endpunkt 'Auftreten eines Lokalrezidivs' Stellung (siehe 5. 2.2). Die Studien wurden deshalb alle ausgeschlossen.

2. Zum Auftreten von Lokalrezidiven nach minimal-invasiver Biopsie wurden neben der Studie von Stolier et al. sieben weitere retrospektive Studien identifiziert (siehe 5.2.3). Es handelte sich um eine Kohortenstudie, fünf Fall-Kontroll-Studien und eine Fallserie. Bei den retrospektiven Fall-Kontrollstudien wurden Patientinnen mit und ohne minimal-invasive Biopsie hinsichtlich des Auftretens von Lokalrezidiven verglichen. Eine Untersuchung der Intervention 'Stichkanalexzision' in Bezug auf das Auftreten der Lokalrezidive erfolgte in keiner Studie. Die Studien beantworten insofern die Frage des Nutzens der Stichkanalexzision nicht und wurden abgeschlossen.

b) Inhaltliche Ergebnisse

Nachweis von Tumorzellverschleppung nach minimal-invasiver Mammabiopsie

Stolier et al. exzidierten nicht durchgängig den gesamten Stichkanal, sondern vor allem Cutis und Subcutis. Der Prozentsatz von im exzidierten Stichkanal nachgewiesenen Tumorzellen nach stereotaktischer minimal invasiver Biopsie lag lediglich bei 2% und betraf zwei Patientinnen mit Mastektomie, bei denen der gesamte Stichkanal exzidiert worden war. Ein Mal war die Exzision ‚en bloc‘ mit dem Operationspräparat erfolgt und ein Mal gesondert davon.

Nutzen der Stichkanalexzision

Beide Patientinnen waren nicht adjuvant bestrahlt worden. Nach 34 Monaten erlitt die Patientin mit gesondert exzidiertem Stichkanal ein dermales Rezidiv in der Stichkanalnarbe.

4.4 Diskussion

a) Methodische Diskussion

Für die vorliegende Fragestellung 'Erforderlichkeit der Stichkanalexzision aufgrund möglicher metastatischer Tumorzellverschleppung nach minimal-invasiver Biopsie' konnte nur eine prospektive Studie identifiziert werden. Weitere identifizierte Studien berichteten zwar von Tumorzellverschleppung, es fehlt aber der Bezug zum klinischen Endpunkt 'Auftreten eines Lokalrezidivs'. Studien mit diesem Endpunkt wurden ebenfalls identifiziert, sie berichteten aber nicht über die Intervention Stichkanalexzision und den Nachweis verschleppter Tumorzellen im Stichkanal.

b) Inhaltliche Diskussion

Für den Nutzen der Stichkanalexzision kann aufgrund unzureichender Datenlage keine Aussage getroffen werden. Es gibt demzufolge bisher keine Evidenz für die Empfehlung zur Exzision des Stichkanals. Allerdings kann auch nicht ausgeschlossen werden, dass es aufgrund einer fehlenden Exzision des Stichkanals zu lokalen Rezidiven kommt, wie sie in Einzelfällen (siehe 4.5) beschrieben sind.

4.5 Studien zu Tumorzellverschleppung nach minimal invasiver Biopsie

Tumorzellverschleppung nach minimal-invasiver Biopsie wurde in jeder der dazu aufgelegten sieben Studien und in drei Fallberichten nachgewiesen, lediglich Liberman et al., 1998, fanden nach Biopsie eines in-situ Karzinoms nur benigne Zellen im Stichkanal (siehe 5.2.2). Der Prozentsatz im Stichkanal nachgewiesener Tumorzellen lag in den Studien zwischen 2,2% und 53,8% bei einem mittleren Zeitraum zwischen Biopsie und Operation von 19,8 Tagen (0 bis 117 Tage). Bei den verwendeten Biopsiemethoden handelte es sich um Stanzbiopsien mit einer Nadelstärke von 14 und 18 Gauge und um Vakuumbiopsien mit einer Nadelstärke von 11 und 14 Gauge.

Am aussagekräftigsten ist die Studie von Diaz et al., 1999. In 325 von 352 konsekutiven minimal-invasiven Biopsien (92,3%) konnte am definitiven Operationspräparat ein Stichkanal nachgewiesen werden. Histologisch wurde der Nachweis durch das Vorhandensein von fibroblastischem und Granulationsgewebe, Fettnekrosen, Zelldetritus, Fremdkörper-Riesenzellreaktionen oder organisierter Hämorrhagie gestellt. In insgesamt 32% wurde nach genau definierten Malignitätskriterien eine Tumorzellverschleppung diagnostiziert.

Der Anteil an verschleppten Tumorzellen war abhängig von der Tumorhistologie und war am größten bei invasiv-duktalem Tumoren und seltener bei in-situ-Karzinom oder lobulärem Karzinom. Der Anteil an Tumorzellverschleppung nahm bei längerem Zeitintervall zwischen Biopsie und definitiver Operation von

42% auf 15% ab. (Die jeweilige Größe der jeweils betrachteten Gruppen zu den beiden vorstehenden Aussagen reicht allerdings nicht aus, um eine statistisch valide Aussage zu treffen). Hinsichtlich der Biopsietechnik und dem Anteil an verschleppten Tumorzellen konnten Diaz et al. zeigen, dass der Anteil von Tumorzellverschleppung bei der Vacuumbiopsietechnik am geringsten ist.

Eine mögliche Tumorzellverschleppung nach minimal-invasiver Biopsie kann aufgrund der vorliegenden Studien als gesichert angenommen werden. Aussagen zur klinischen Bedeutung der Tumorzellverschleppung sind aufgrund dieser Studien nicht möglich.

4.6 Studien zum Auftreten von Lokalrezidiv nach minimal-invasiver Biopsie

Neben der Studie von Stolier et al. zeigt eine retrospektive Fallserie (Uriburu et al., 2006) aufgrund der Lokalisation von drei Lokalrezidiven dermal oder subdermal in der Stichkanalnarbe nach hautsparender Mastektomie einen klinisch plausiblen Zusammenhang zu einer Tumorzellverschleppung bei der Stanzbiopsie.

Drei Einzelfallberichte zum Auftreten von Lokalrezidiven (Chao et al., 2001; Intra et al., 2005 und Kwo et al., 2006) zeigen ebenfalls einen klinisch plausiblen Zusammenhang mit der minimal-invasiven Biopsie bei dermal in der Stichkanalnarbe oder pectoral nach Biopsie eines in-situ-Karzinoms auftretenden Rezidiven.

Thurfjell et al., 2000 berichtet von drei aufgrund Tumorzellverschleppung vermuteter intramammärer Rezidive, die jeweils als konkordant mit der stanzbiopsischen Nadelführung 6-10 cm vom Tumorbett entfernt beschrieben werden. Hier erscheint der Zusammenhang zur minimal-invasiven Biopsie fraglicher, da die Rezidive nicht durch ein eindeutiges Korrelat zugeordnet werden können.

Bezüglich der klinischen Relevanz der verschleppten Tumorzellen im Hinblick auf das Auftreten eines Lokalrezidivs konnten also insgesamt acht Fälle (aus fünf Publikationen) identifiziert werden, in denen ein klinisch sehr plausibler Zusammenhang zwischen minimal-invasiver Biopsie und Lokalrezidiven aufgrund von verschleppten Tumorzellen aufgezeigt wurde.

Die beschriebenen Lokalrezidive traten alle bei Patientinnen auf, die nicht adjuvant bestrahlt wurden. Eine Eliminierung verschleppter Tumorzellen durch die Radiatio erscheint klinisch plausibel.

Die weiterhin vorliegenden Fall-Kontroll-Studien zum Auftreten von Lokalrezidiven ohne und mit minimal-invasiver Biopsie zeigen insgesamt – bei schlechter methodischer Qualität – keine erhöhte Lokalrezidivrate nach minimal-invasiver Biopsie im Vergleich zur offenen Biopsie. In Bezug auf den Vergleich von Rezidivraten sind die Fallzahlen dieser Fall-Kontrollstudien zu gering, um statistisch signifikante Vergleichszahlen zu erbringen. Bei zwei Studien war die Vergleichbarkeit von Fall- und Kontrollgruppen darüberhinaus als eingeschränkt zu werten.

5. Evidenztabellen

5.1 Evidenz von Einzelstudien

5.1.1 Eingeschlossene Einzelstudien

Tabelle 1

Evidenz von Einzelstudien zum Nutzen der Stichkanalexzision nach minimal-invasiver Biopsie

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung	Evidenz- grad
Stolier et al., 2000, USA	<p>Outcome-Studie</p> <p>n= 89, prospektiv alle mit minimal-invasiver Biopsie, unterschiedliche Methoden (stereotaktisch 'multiple- puncture' CNB 14g, stereo-taktisch vacuum-assistiert 11g oder 14 g, Sonogest. 14g)</p> <p>Bei OP Exzision von Einstichstelle zumindest Cutis und Subcutis.</p> <p>15,7 %Tis 62,9 %T1 16,9 %T2 4,5 %T3</p> <p>Follow-up: mind. 34 Monate</p>	<p><u>Outcome-Parameter 1:</u> Tumordinfiltration von Cutis oder Subkutis des Stichkanals (kein Bezug zu Primärtumor)</p> <p>n= 2 (11 und 13 Tage nach Biopsie) = 2,2% (jeweils T1c)</p> <p><u>Operationsverfahren:</u> <i>Beide Mastektomie ohne Radiatio, Exzision des gesamten Stichkanals 1x gesondert, 1x im Mastektomiepräparat enthalten</i></p> <p><u>Outcome-Parameter 2:</u> Rezidiv n=1 an der Biopsie-einstichstelle nach 34 Monaten bei Patientin mit Exzision des Stichkanals</p>	Beide Patientinnen mit Tumordinfiltration der Haut hatten multiple Einstiche (stereo-taktisches Verfahren mit 14g Nadel) d.h. Hautinfiltration mit dieser Technik war 25% (2/8), für die anderen Biopsiemethoden 0%.	<p>Einstichstelle soll bei Mastektomie so entfernt werden, dass sie ‚en bloc‘ bleibt, auch wenn die Resektionsgrenzen dadurch verändert werden, es sei denn, postoperative Radiatio ist geplant</p> <p>Fallzahl lässt keine sichere Schlussfolgerung über Risiko des einzelnen Punktionsverfahrens zu!</p>	2c

5.2 Ausgeschlossene Studien

5.2.1 Ausgeschlossene aggregierte Evidenz

Tabelle 2: Ausgeschlossene Reviews zum Nachweis von Tumorzellverschleppung nach minimal-invasiver Biopsie

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Bellocq et al., 1998, Frankreich	Review versch. Aspekten der Stanzbiopsie u.a. Tumorzellverschleppung	Narrativ
Janssens et al., Belgien, 2006	Review zur 'Ballistik' verschleppter Tumorzellen und Biopsietechniken	Narrativ
Liberman, 2000, USA	Review zu versch. Aspekten der Stanzbiopsie u.a. Tumorzellverschleppung mit weitgehend vollst. Auflistung von Studien bis 2000	Nur Arbeiten aus USA berücksichtigt
Smith, 1984, USA	Survey zu FNAC und Tumorzellverschleppung bei versch. Tumorentitäten	FNAC, Brustkrebs nicht eingeschl.
Tardivon et. al., 2002, Frankreich	Reviews zu verschiedenen Aspekten der Stanzbiopsie u.a. Tumorzellverschleppung	Narrativ
Thurfjell, 2001, Schweden	Review zu versch. Aspekten des Mammographie-Screening u.a. Tumorzellverschleppung	v.a. Thurfjells eigene Arbeit (s. Tabelle 3) wird diskutiert
Usami et al., 2005, Japan	Review zu versch. Aspekten der Stanzbiopsie u.a. Tumorzellverschleppung	Narrativ

5.2.2 Studien zum Nachweis von Tumorzellverschleppung

Tabelle 3
Studien mit Nachweis von Tumorzellverschleppung im exzidierten Stichkanal

Autor, Jahr, Land	n Gesamt	Biopsiemethode	Prozentsatz verschleppter Tumorzellen	Zeitraum Biopsie-OP	Bemerkung
Diaz et al., 1999, USA	352 (1995-1998, konsekutiv)	CNB 14g und 18g VAB 11g und 14g gesamt :	32% (114/352)	2-117 Tage, (med. 26)	Bei längerem Zeitintervall wurde weniger Tumorzellverschleppung nachgewiesen: 0-14 Tage (n= 158): 42% 15-28 Tage (n= 121): 31% >28 Tage (n= 73): 15% in 7,7% kein Stichkanalnachweis (n= 25)
Hoorntje et al., 2004, Niederlande	(a) 64	CNB 14g retrospektiv	17,2 % (11/64)	0, OP immer am gleichen Tag	34,3% kein Stichkanalnachweis (n=22) 15,3% kein Stichkanalnachweis (n=2)
	(b) 13	CNB 14g prospektiv	53,8% (7/13)	7-35Tage (med. 21)	Verschleppte Tumorzellen innerhalb und außerhalb des Tumors wurden berücksichtigt

Autor, Jahr, Land	n Gesamt	Biopsiemethode	Prozentsatz verschleppter Tumorzellen	Zeitraum Biopsie-OP	Bemerkung
Liberman et al., 1998, USA	28 (DCIS)	VAB 11g	7% (2/28)	10-59 (med. 27)	Nur benigne Zellen!
Stolier et al., 2000, USA	89	stereotaktisch 'multiple- puncture' CNB 14g, stereotaktisch vacuum-assistiert 11g oder 14g, US-gest. 14g)	2,2% (2/89)	-	1 Lokalrezidiv trotz exzidierten Stichkanals (s.u.)
Youngson, et al., 1994	43	CNB 14g	28% (12/43)	0-112 (med. 25)	Neben CNB wurde auch Lokalanästhesie appliziert oder es fand eine Drahtmarkierung statt (von Autor als mögliche Confunder gesehen)

Tabelle 4

Fallberichte zu Tumorzellverschleppung nach minimal-invasiver Biopsie

Autor, Jahr, Land	Inhalt
Douglas-Jones et al., 2002, UK	(1 von 2 Fällen beschrieben): 3 Wochen nach 14g CNB zeigen sich im Tumorpräparat neben einem intrazystischen papillären Karzinom ohne Wandinfiltration in Richtung Haut papilläre Tumorzellnester in einer Fettnekrosezone vereinbar mit dem Stichkanal
Harter et al., 1998, USA	2 Wochen nach stereotaktischer CNB (14g) zeigen sich im Tumorexzisionspräparat (Wide excision) 2 cm vom Tumorrand zwischen Tumor und Haut –distinkt. von der WL- zahlreiche Zellkluster des muzinösen Karzinoms, umgeben von hämosiderin-beladenen Makrophagen als Indiz für die Platzierung im Stichkanal

Tabelle 4

Nicht berücksichtigte Studien zu Tumorzellverschleppung nach minimal-invasiver Biopsie

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Boppana et al., 1994, USA	FNAC mit 22g, Tumorzellverschleppung in 36/100 Fällen (36%)	FNAC und nur als abstract publiziert
Bott et al., 1998, UK	4 Wochen nach FNAC zeigt Mastektomiepräparat den FNAC-Stichkanal und darin vitale Zellen des infiltrativ-duktales Karzinoms	FNAC
Nagi et al., 2005, USA	Review von 53 Fällen mit verschleppten Epithelzellen. In Bezug auf durchgeführte Nadelprozeduren In 51 Fällen war eine papilläre Läsion nachweisbar (maligne oder benigne)	Es kann keine Aussage über Anteil an Tumorzellverschleppung eines Gesamtkollektivs im Hinblick auf minimal-invasive Biopsieverfahren gemacht werden
Youngson, 1994, USA	31 Fälle mit disloziertem Epithel bei Brusttumoren (maligne und benigne) werden in Bezug auf durchgeführte Nadelprozeduren analysiert	Es kann keine Aussage über Anteil an Tumorzellverschleppung eines Gesamtkollektivs im Hinblick auf minimal-invasive Biopsieverfahren gemacht werden

5.2.3 Studien mit dem Endpunkt Rezidiv nach minimal-invasiver Biopsie

Tabelle 5

Studien mit dem Endpunkt Rezidiv nach minimal-invasiver Biopsie

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung
Chen et al., 2002, USA	<p>Retrospektive Fall-Kontroll-Studie</p> <p>n= 551, alle konsekutiv mit BET und Radiatio 7/92-12/98</p> <p>(a) n= 86 mit CNB (11g und 14 g stereotaktisch)</p> <p>(b) n= 242 mit WL</p> <p>(c) n= 223 ohne IG</p> <p>(a) und (b) mit vergleichbaren Tumorgößen und vergleichbarer adjuvanter Therapie, (c) mit signifikant größeren Tumoren</p> <p>(a) hatte häufiger R1-Resektion (6,9%) als (b) und (c) (jeweils unter 2%)</p> <p>medianes Follow-Up 4,9 Jahre (2,0-8,9)</p>	<p>Lokalrezidivrate:</p> <p>(a) 2,3% (2/86)</p> <p>(b) 5,4% (13/252)</p> <p>(c) 10,3% (23/223)</p> <p>Rezidive bei (a):</p> <ol style="list-style-type: none"> 1. andere Histologie und anderer Quadrant als Primärtumor (eher Zweittumor) 2. R1-Resektion für umgebendes DCIS bei invasivem Karzinom, Rezidiv zeigt DCIS an der Stelle der R1-Resektion 	<p>(a) und (b) vergleichbar</p> <p>11-14g stereotaktische Biopsien</p> <p>Exzision des Stichkanals wurde nicht extra durchgeführt</p> <p>Die Studie ist statistisch bei n=86 und Rezidivrate 2,3% nicht ausreichend gepowert, um den Unterschied zwischen a) und b) als statistisch signifikant auszuweisen</p>	<p>Vor allem durch adjuvante Radiatio werden Metastasen durch NTS verhindert</p>

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung
Fitzal et al., 2006, Österreich	<p>Retrospektive Fall-Kontroll-Studie</p> <p>n= 719 (a) n=189 mit CNB (b) n=530 ohne CNB</p> <p>alle mit BET, R0 Resektion (mindestens 1 mm)</p> <p>100% adjuvante Radiatio</p> <p>Gruppen hinsichtlich T-Stadien, Alter, Grading und ER-Status, sowie Systemtherapie vergleichbar, nur PgR-Status in Biopsie-Gruppe signifikant häufiger+.</p> <p>Medianes Follow-up: (a)= 78 Monate (b)= 71 Monate</p>	<p>Lokalrezidivrate:</p> <p>(a) 1,1% (b) 2,1%</p>	<p>Es wurde nicht erhoben, wie häufig eine Stichkanalexzision durchgeführt wurde</p> <p>Autoren geben an, dass die Exzision grundsätzlich das empfohlene Vorgehen war</p> <p>Rezidive nicht nach Topographie in Bezug auf CNB beschrieben</p> <p>Gruppen sehr unterschiedlich groß, auch diese Fallzahl reicht nicht für statistisch signifikante Aussage</p>	<p>CNB zeigt keine höhere Rezidivrate bei adjuvanter Radiatio als bei offener Biopsie</p> <p>Empfehlen Vermeidung von 14g oder < und Vermeidung von ‚multiple-puncture‘-Technik</p>

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung
King et al., 2001, USA	<p>Retrospektive Fall-Kontroll-Studie</p> <p>CNB (stereotaktisch oder US-gesteuert, 14g) versus Drahtmarkierung (20g)</p> <p>n= 211 von 351 konsekutiv mit BET behandelten Pat. mit nichtpalpablen Tumoren 7/93-7/06</p> <p>(a) = 132 CNB (b) = 79 WL</p> <p>Gruppen mit vergleichbaren Histologien und adjuvanten Therapien</p> <p>alle mit tumorfreien Rändern</p> <p>Radiatio in 91%</p> <p>Medianes Follow-up: (a) = 44,4 Monate (b) = 50,1 Monate</p>	<p>Rezidivrate:</p> <p>(a) = 3% (4) (b) = 2,5% (2)</p> <p>keine statistisch signifikante Differenz</p>	<p>Rezidive nicht topographisch in Bezug auf CNB oder WL Stichkanal beschrieben</p> <p>Keine Exzision des Stichkanals durchgeführt</p> <p>Die Studie ist bei n= 132/79 für eine statistisch signifikante Aussage nicht ausreichend gepowert</p>	<p>Empfehlen nicht, den Stichkanal zu exzidieren, adjuvante Radiatio nach BET wird als ausreichend angesehen</p>

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung
Knight et al., 2002, USA	<p>Retrospektive Fall-Kontroll-Studie</p> <p>n= 398, alle im Zeitraum mit BET und CNB oder WL von 5/90-6/95</p> <p>(a) n= 297 mit CNB (14g) (b) n= 101 mit WL</p> <p>94% mit Radiatio</p> <p>Keine vergleichenden Angaben zu R-Status, T-Stadienverteilung, Grading, Rezeptorstatus oder Systemtherapie</p> <p>Mittleres/Medianes? Follow-up: 29,7 Monate (2-90)</p>	<p>Lokalrezidivrate:</p> <p>(a) 3,7% (11) (b) 3,9% (4)</p> <p>nicht statistisch signifikant unterschiedlich</p>	<p>Exzision des Stichkanals wurde nicht extra durchgeführt</p> <p>Rezidivrate bei Patientinnen ohne adjuvante Radiatio (n=24) 12,5%</p> <p>Auch hier reicht die Fallzahl nicht, um eine statistische Signifikanz zu berechnen</p>	<p>CNB zeigt im Vergleich zu WL bei offener Biopsie keine höhere Rezidivrate</p> <p>Entscheidend ist adjuvante Radiatio</p>
Thurfjell et al., 2000, Schweden	<p>Retrospektive Kohortenstudie</p> <p>n= 303 konsekutive Patientinnen mit nichtpalpablen Tumoren und BET</p> <p>CNB/FNA in 71% (214/303)</p> <p>Radiatio in 84,5%</p> <p>Follow-up: 9 Jahre</p>	<p>Anzahl (vermuteter) Rezidive aufgrund der CNB/FNAC: 9% (3/33)</p> <p>Angabe von Gründen:</p> <ol style="list-style-type: none"> 1. gleiche Histologie 2. alle mindestens 5mm tumorfrei (pT1b und pT1c) 3. alle Rezidive 6-10 cm vom Tumorbett entfernt 4. Lokalisation jeweils als konkordant mit präoperativer Nadelführung beschrieben, 2x jedoch auch mit Drahtmarkierung 	<p>Gesamtzahl der Rezidive nicht stratifiziert nach Patientinnen mit und ohne CNB</p> <p>die 3 Patientinnen mit vermutetem Rezidiv nach minimal-invasiver Diagnosesicherung hatten alle keine adjuvante Radiatio</p> <p>Alle Patientinnen hatten Drahtmarkierung</p> <p>Exzision des Stichkanals nicht beschrieben</p>	<p>Lokalrezidivrisiko aufgrund NTS ist v.a. bei fehlender adjuvanter Radiatio erhöht</p> <p>Empfehlen Exzision des Stichkanals bei CNB/FNA und Platzierung der Drahtmarkierung dorsal des Tumors</p>

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Empfehlungen/ Schlussfolgerung
Uriburu et al., 2006, Argentinien	<p>Retrospektive Fallserie</p> <p>n= 58 im Zeitraum 12/99-12/02 mit SSM und Sofortrekonstruktion (TRAM). davon n=11 mit Biopsie (VAB/CNB)</p> <p>alle R0</p> <p>keine adjuvante Radiatio</p> <p>med. Follow-up: 28 Monate (5-60 Monate)</p>	<p>Anzahl (vermuteter) Rezidive nach SLCB: 3/3 (nach 16, 22 und 23 Monaten) alle mit 14g CNB, stereotaktisch, jeweils mind. 3 Punktionen bei multizentrischem Karzinom</p> <p>Angabe von Gründen:</p> <p>1. alle 3 Rezidive traten dermal oder subdermal an der Biopsienarbe auf, nicht an der Mastektomienarbe</p> <p>2. jeweils gleiche Histologie wie Primärtumor</p>		<p>Exzision des Stichkanals mit umgebendem Hautmantel wird empfohlen</p>

5.2.4 Fallberichte mit dem Endpunkt Rezidiv nach minimal-invasiver Biopsie

Tabelle 6

Fallberichte zum Auftreten von Rezidiv nach minimal-invasiver Biopsie

Autor, Jahr, Land	Inhalt
Chao et al., 2001, USA	Intraoperativ, sowie nach 13 und 17 Monaten werden 3 subkutane Rezidive (1x dermale Mikrometastasen) an der Biopsie-Einstichstelle bei 3 Patientinnen mit Biopsie (multiple Punktionen mit 14g) und modifiziert radikaler Mastektomie diagnostiziert, jeweils keine adjuvante Radiatio
Intra et al., 2005, Italien	2 Jahre nach Mastektomie wegen DCIS-Rezidiv (G3 mit Nekrosen), diagnostiziert mit stereotaktischer CNB, zeigt sich ein 1cm großes infiltratives intrapektorales Rezidiv 2 cm oberhalb der Mastekomienarbe, keine adjuvante Radiatio
Kwo et al., 2006, USA	Patient mit CNB bei 12.00 und Skin Sparing Mastektomie mit TRAM-Sofortaufbau zeigt nach 6 Jahren genau in der CNB-Einstichstelle eine Hautmetastase, keine adjuvante Radiatio

Tabelle 7

Nicht berücksichtigte Studien zum Rezidiv nach Feinnadelpunktion

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Berg et al., 1961, USA	Retrospektive Fall-Kontroll-Studie (matched-pairs) n= 370 Fälle mit Aspirationsbiopsie g=? n= 370 Kontrollen gematcht nach Alter (5 Jahresgruppen), histologischem Typ, Grading, Größe des Primärtumors und LK-Status von 1.412 operierten Pat. mit postoperativer Histologie im Zeitraum 1940-43 Die 15-Jahres-Überlebensrate ist für die Pat. mit Biopsie mind. gleich gut wie die für Pat. ohne Biopsie.	FNAC keine (Lokal)-Rezidivrate angegeben, Anzahl von ‚lost of Follow-up‘ unklar
Robbin et al., 1954	Retrospektive Fall-Kontroll-Studie (matched-pairs) gleiche Studie wie Berg et al. mit Publikation erster Ergebnisse	FNAC Studie liegt noch nicht im Original vor
Taxin et al., 1997, USA	Retrospektive Fall-Kontroll-Studie n=58 FNAC n=308 Kontrollen Vergleichbar in Bezug auf TNM, Histologie, Grading, Rezeptorenstatus und operativer sowie adjuvanter Therapie FNAC Gruppe im Mittel 8 Jahre jünger 5-Jahres Lokalrezidiv- und Fernmetastasen freies Überleben FNAC-Gruppe: 94% und 86% Kontrollen: 86% und 85% FNAC nicht signifikant für Rezidiv	FNAC, Studie für Statistik nicht mit ausreichender Fallzahl

6. Literatur

6.1 Aufbereitete Evidenz - ausgeschlossene Reviews

1. Bellocq JP, Zafrani B, Chenard MP. [Diagnostic difficulties and limits in breast histopathology in core biopsies (breast microbiopsies)]. *Arch Anat Cytol Pathol* 1998;46(4):257-60
2. Janssens JP, Rotenberg L, Sentis M, Motmans K, Schulz-Wendtland R. Caution with microbiopsies of the breast: displaced cancer cells and ballistics. *Eur J Cancer Prev* 2006;15(6):471-3
3. Liberman L. Clinical management issues in percutaneous core breast biopsy. *Radiol Clin North Am* 2000;38(4):791-807
4. Smith EH. The hazards of fine-needle aspiration biopsy. *Ultrasound Med Biol* 1984;10(5):629-34
5. Tardivon AA, Guinebretiere JM, Dromain C, Deghaye M, Caillet H, Georgin V. Histological findings in surgical specimens after core biopsy of the breast. *Eur J Radiol* 2002;42(1):40-51
6. Thurffjell MG. Aspects in mammographic screening. Detection, prediction, recurrence and prognosis. *Acta Radiol Suppl* 2001;42(424):1-22
7. Usami S, Moriya T, Kasajima A, Suzuki A, Ishida T, Sasano H, Ohuchi N. Pathological aspects of core needle biopsy for non-palpable breast lesions. *Breast Cancer* 2005;12(4):272-8

6.2 Evidenz von Einzelstudien zum Nutzen der Stichkanalexzision

1. Stolier A, Skinner J, Levine EA. A prospective study of seeding of the skin after core biopsy of the breast. *Am J Surg* 2000;180(2):104-7

6.3 Studien und Fallberichte zu Tumorzellverschleppung nach minimal-invasiver Intervention

1. Boppana S, May M, Hoda S. Does prior fine-needle aspiration cause diagnostic difficulties in histologic evaluation of breast carcinomas? (abstr.). *Lab Invest* 1994;70:13A
2. Bott S, Mohsen V, Wells C, Carpenter R. Needle tract metastases in breast cancer. *Eur J Surg Oncol* 1999;25(5):553
3. Diaz LK, Wiley EL, Venta LA. Are malignant cells displaced by large-gauge needle core biopsy of the breast? *AJR Am J Roentgenol* 1999;173(5):1303-13
4. Douglas-Jones AG, Verghese A. Diagnostic difficulty arising from displaced epithelium after core biopsy in intracystic papillary lesions of the breast. *J Clin Pathol* 2002;55(10):780-3
5. Harter LP, Curtis JS, Ponto G, Craig PH. Malignant seeding of the needle track during stereotaxic core needle breast biopsy. *Radiology* 1992;185(3):713-4
6. Hoorntje LE, Schipper ME, Kaya A, Verkooijen HM, Klinkenbijn JG, Borel R, I. Tumour cell displacement after 14G breast biopsy. *Eur J Surg Oncol* 2004;30(5):520-5
7. Liberman L, Smolkin JH, Dershaw DD, Morris EA, Abramson AF, Rosen PP. Calcification retrieval at stereotactic, 11-gauge, directional, vacuum-assisted breast biopsy. *Radiology* 1998;208(1):251-60
8. Nagi C, Bleiweiss I, Jaffer S. Epithelial displacement in breast lesions: a papillary phenomenon. *Arch Pathol Lab Med* 2005;129(11):1465-9

9. Youngson BJ, Liberman L, Rosen PP. Displacement of carcinomatous epithelium in surgical breast specimens following stereotaxic core biopsy. *Am J Clin Pathol* 1995;103(5):598-602
10. Youngson BJ, Cranor M, Rosen PP. Epithelial displacement in surgical breast specimens following needling procedures. *Am J Surg Pathol* 1994;18(9):896-903

6.4 Studien- und Fallberichte mit dem Endpunkt Lokalrezidiv nach minimal-invasiver Intervention (einschließlich Feinnadelpunktion)

1. Berg JW, Robbins GF. A late look at the safety of aspiration biopsy. *Cancer* 1962;15:826-7
2. Chao C, Torosian MH, Boraas MC, Sigurdson ER, Hoffman JP, Eisenberg BL, Fowble B. Local recurrence of breast cancer in the stereotactic core needle biopsy site: case reports and review of the literature. *Breast J* 2001;7(2):124-7
3. Chen AM, Haffty BG, Lee CH. Local recurrence of breast cancer after breast conservation therapy in patients examined by means of stereotactic core-needle biopsy. *Radiology* 2002;225(3):707-12
4. Fitzal F, Sporn EP, Draxler W, Mittlbock M, Taucher S, Rudas M, Riedl O, Helbich TH, Jakesz R, Gnant M. Preoperative core needle biopsy does not increase local recurrence rate in breast cancer patients. *Breast Cancer Res Treat* 2006;97(1):9-15
5. Intra M, Mazzarol G, Rietjens M, az Brito JA, Gennari R, Soteldo J, Rodriguez J, Bassani G, Bassi F. Extramammary recurrence of DCIS after total mastectomy: an iatrogenic displacement following needling procedures? *Breast J* 2005;11(4):297-300
6. King TA, Hayes DH, Cederbom GJ, Champaign JL, Smetherman DH, Farr GH, Bolton JS, Fuhrman GM. Biopsy technique has no impact on local recurrence after breast-conserving therapy. *Breast J* 2001;7(1):19-24
7. Knight R, Horiuchi K, Parker SH, Ratzer ER, Fenoglio ME. Risk of needle-track seeding after diagnostic image-guided core needle biopsy in breast cancer. *JSLs* 2002;6(3):207-9
8. Kwo S, Grotting JC. Does stereotactic core needle biopsy increase the risk of local recurrence of invasive breast cancer? *Breast J* 2006;12(3):191-3
9. Moore KH, Thaler HT, Tan LK, Borgen PI, Cody HS, III. Immunohistochemically detected tumor cells in the sentinel lymph nodes of patients with breast carcinoma: biologic metastasis or procedural artifact? *Cancer* 2004;100(5):929-34
10. Robbins GF, Brothers H, Eberhart WF, Quan S. Is aspiration biopsy of breast cancer dangerous to the patient? *Cancer* 1954;7(4):774-8
11. Taxin A, Tartter PI, Zappetti D. Breast cancer diagnosis by fine needle aspiration and excisional biopsy. Recurrence and survival. *Acta Cytol* 1997;41(2):302-6
12. Thurfjell MG, Jansson T, Nordgren H, Bergh J, Lindgren A, Thurfjell E. Local breast cancer recurrence caused by mammographically guided punctures. *Acta Radiol* 2000;41(5):435-40
13. Uriburu JL, Vuoto HD, Cogorno L, Isetta JA, Candas G, Imach GC, Bernabo OL. Local recurrence of breast cancer after skin-sparing mastectomy following core needle biopsy: case reports and review of the literature. *Breast J* 2006;12(3):194-8

A Mammographie-Screening
bei Frauen 40-49 Jahre und >70 Jahre
Nutzen/Schaden-Bilanz

B Doppelbefundung von Screening-Mammographien
unter Berücksichtigung des
Entscheidungsverfahrens bei diskrepanten
Befunden und des Einsatzes von Computer-
Assistierten-Diagnose (CAD)-Systemen

C Stellenwert der Mammasonographie im Rahmen
der Brustkrebsfrüherkennung

D Indikationsstellung für die stereotaktische
Vakuumbiopsie im Vergleich zur stereotaktischen
Stanzbiopsie

E Nutzen der Exzision des Stichkanals nach minimal-
invasiver Biopsie der Brust zur Vermeidung von
Tumorzellverschleppung

F **Her-2/neu-Bestimmung:
Vergleich der Methoden
Immunhistochemie,
Fluoreszenz-in-situ-Hybridisierung und
Chromogene-in-situ-Hybridisierung**

Inhalt

1. Hintergrund	194
2. Fragestellung	197
3. Methoden	197
3.1 Kriterien für den Einschluss- und Ausschluss von Studien	197
3.1.1 Zielgruppe	197
3.1.2 Diagnostische Verfahren	197
3.1.3 Studientypen	198
3.1.4 Erforderliche Angaben	198
3.1.5 Einschluss- und Ausschlusskriterien	198
3.2 Literaturrecherche	198
3.3 Zielgrößen und deren Definition	199
3.4 Bewertung der gefundenen Literatur	200
4. Ergebnisse	201
4.1 Vergleich der HER-2/neu-Bestimmung durch IHC und FISH in Bezug auf das Ansprechen auf die Therapie mit Trastuzumab	201
4.1.1 Ergebnisse der Literaturrecherche	201
4.1.2 Aufbereitete Evidenz	201
4.1.3 Evidenz von Einzelstudien	201
4.1.4 Diskussion	202
4.2 Reliabilität der Methoden IHC und FISH	203
4.2.1 Ergebnisse der Literaturrecherche	203
4.2.2 Aufbereitete Evidenz	203
4.2.3 Evidenz von Einzelstudien	203
4.2.4 Diskussion	204
4.3 Konkordanz von IHC und FISH	204
4.3.1 Ergebnisse der Literaturrecherche	204
4.3.2 Aufbereitete Evidenz	205
4.3.3 Evidenz von Einzelstudien	205
4.3.4 Diskussion	206
4.4 Konkordanz und Reliabilität von FISH und CISH	207
4.4.1 Ergebnisse der Literaturrecherche	207
4.4.2 Aufbereitete Evidenz	207
4.4.3 Evidenz von Einzelstudien	207
4.4.4 Diskussion	208
4.5 Konkordanz der Her2/neu-Bestimmung aus Biopsiegewebe im Vergleich mit dem definitiven Operationspräparat	208
4.5.1 Ergebnisse der Literaturrecherche	208
4.5.2 Aufbereitete Evidenz	208
4.5.3 Evidenz von Einzelstudien	208
4.5.4 Diskussion	209

5.	Evidenztabellen	210
5.1	Ergebnisse und Evidenzgrad von IHC und FISH in Bezug auf das Ansprechen auf Trastuzumab.....	210
5.1.1	Eingeschlossene Studien	210
5.1.2	Ausgeschlossene Studien	213
5.2.	Ergebnisse und Evidenzgrad zur Reliabilität von IHC und FISH.....	214
5.2.1	Ausgeschlossene Studien zur Reliabilität von IHC und FISH.....	224
5.3.	Ergebnisse und Evidenzgrad Konkordanz IHC und FISH.....	225
5.3.1	Ausgeschlossene Studien zur Konkordanz IHC und FISH.....	246
5.4.	Konkordanz von FISH und CISH	248
5.4.1	Ergebnisse und Evidenzgrad FISH-CISH	248
5.5	Her-2/neu-Bestimmung im Vergleich Stanzbiopie-definitives Operationspräparat	258
5.5.1	Ergebnisse und Evidenzgrad Her-2/neu-Bestimmung im Vergleich Stanzbiopsie – definitives Operationspräparat.....	258
5.5.2	Ausgeschlossene Studien im Vergleich Her-2/neu aus Stanzbiopsie - definitives Operationspräparat	263
6.	Literatur.....	264

Tabellenverzeichnis

Tabelle 1	Ergebnisse und Evidenzgrad von IHC und FISH zum Therapieansprechen auf Trastuzumab	210
Tabelle 2	Ausgeschlossene Studien zum Therapieansprechen auf Trastuzumab	213
Tabelle 3	Ergebnisse und Evidenzgrad zur Reliabilität von IHC und FISH	214
Tabelle 4	Ausgeschlossene Studien zur Reliabilität von IHC und FISH	224
Tabelle 5	Ergebnisse und Evidenzgrad Konkordanz IHC und FISH -I- stratifiziert nach 0/1+, 2+, 3+	225
Tabelle 6	Ergebnisse und Evidenzgrad Konkordanz IHC und FISH -II- nur nach +/- stratifiziert oder nur für IHC2+ oder IHC 3+	237
Tabelle 7	Ausgeschlossene Studien zur Konkordanz IHC-FISH	246
Tabelle 8	Ergebnisse und Evidenzgrad FISH-CISH	248
Tabelle 9	Her-2/neu Bestimmung im Vergleich Stanzbiopise-definitives Operationspräparat	258
Tabelle 10	Ausgeschlossene Studien im Vergleich Her-2/neu aus Stanzbiopsie - definitives Operationspräparat	263

1. Hintergrund

Epidermale Wachstumsfaktor-Rezeptoren finden sich auf jeder epidermalen Zelloberfläche. Die an diese Rezeptoren bindenden Wachstumsfaktoren nehmen Einfluss auf die Zellteilung [1]. Bei einem Teil maligner solider Tumoren, darunter auch Brustkrebs, lässt sich eine vermehrte Expression von Wachstumsfaktor-Rezeptoren nachweisen, diese geht einher mit einem aggressiveren Tumorwachstum. Im letzten Jahrzehnt wurden spezifische Antikörper gegen Wachstumsfaktor-Rezeptoren entwickelt [2;3].

Bei Überexpression des humanen epidermalen Wachstumsfaktor-Rezeptors Her-2/neu (im Folgenden Her-2/neu) oder bei Genamplifikation des für diesen Wachstumsfaktor-Rezeptor kodierenden Genabschnitts auf Chromosom 17 kann der monoklonale Antikörper Trastuzumab seit Mai 2006 in Europa nicht nur bei metastasiertem Brustkrebs, sondern auch in der adjuvanten Situation therapeutisch eingesetzt werden [3]. Die Trastuzumab-Therapie ist mit zum Teil erheblichen, vor allem kardialen Nebenwirkungen verbunden [4]. Den Bestimmungsmethoden für Her-2/neu kommt deshalb hinsichtlich ihrer diagnostischen Sicherheit größte Bedeutung zu.

Der Her-2/neu-Status wird darüberhinaus seit der Konsensus Konferenz in St. Gallen 2005 als unabhängiger Prognosefaktor gewertet [5] und hat einen prädiktiven Wert hinsichtlich des Ansprechens auf Systemtherapien [6].

Anerkannte Verfahren zur Bestimmung des Her-2/neu-Status sind die Immunhistochemie (IHC) und die in-situ-Hybridisierung [7].

Immunhistochemisch wird die Expression des Her2-neu-Rezeptors an der Zellmembran gemessen. Dieses Verfahren ist analog der Bestimmung für Hormonrezeptoren in jedem histopathologischen Labor möglich.

Die Fluoreszenz-in-situ-Hybridisierung (FISH) weist die Amplifizierung des für den Her-2/neu-Rezeptor kodierenden Gen-Abschnitts auf Chromosom 17 nach. Das Verfahren ist aufwändig und teuer, u.a. da es ein Fluoreszenz-Mikroskop erfordert. Die FISH wird deshalb nur in wenigen zentralen Laboren durchgeführt. Seit einigen Jahren gibt es alternativ zur Fluoreszenz-in-situ-Hybridisierung die Möglichkeit, eine chromogene-in-situ-Hybridisierung (CISH) zum Nachweis der Genamplifikation durchzuführen, die mit einem Lichtmikroskop einfacher durchzuführen und wesentlich kostengünstiger ist als die FISH [8]. Die CISH hat sich in Deutschland bisher nur wenig durchgesetzt.

Derzeitiger Standard der Her2-neu-Bestimmung in Deutschland ist die Rezeptormessung mittels Immunhistochemie mit einem Scoring von 0-3+ (anerkanntes Scoring nach DAKO, dem Hersteller des immunhistochemischen Testkits Herceptest) und die zusätzliche Testung durch FISH bei Vorliegen der Kategorie 2+ [7].

Patientinnen mit der immunhistochemischen Kategorie 3+ gelten als sicher positiv. Für diese Kategorie wird ein Rezeptornachweis durch eine komplette, starke Zellmembrananfärbung bei über 10% der Zellen gefordert. Patientinnen der Kategorien 0 (membranöse Färbung < 10% der Zellen) und 1+ (schwache, kaum sichtbare unvollständige Membrananfärbung in mehr als 10% der Zellen) gelten als sicher negativ im Hinblick auf eine Überexpression des Rezeptors.

Die Kategorie 2+ (eine schwache bis mäßige vollständige Membrananfärbung in über 10% der Zellen) wird als unsicher in Bezug auf eine Überexpression bewertet [8].

Für den Nachweis der Genamplifizierung durch in-situ-Hybridisierung des entsprechenden Genabschnitts galt bisher eine Einteilung in negativ und positiv [9].

Durch die in-situ-Hybridisierung wird die Anzahl der Genkopien für den Wachstumsfaktor-Rezeptor nachgewiesen. Normalerweise sind zwei Genkopien vorhanden. Von einer Amplifizierung geht man bei einer Anzahl von mehr als fünf Genkopien pro Zellkern aus.

Bei Polysomie des Chromosoms 17 liegen ohne Amplifizierung ebenfalls mehrere Genkopien vor. Neuere FISH-Verfahren weisen auch die Anzahl der Centromer 17-Signale nach. Die FISH mit einer Korrektur auf Polysomie gilt dann als positiv, wenn die Anzahl der Genkopien geteilt durch die Anzahl der Centromer 17-Signale größer als zwei ist.

Eine Therapie mit Trastuzumab ist bei Patientinnen indiziert, die Her-2/neu positiv sind. Als positiver Her-2/neu-Status gilt die immunhistochemische Rezeptorbestimmung mit einem Ergebnis von 3+ oder die positive in-situ-Hybridisierung mit einem Ergebnis von > 5 bzw. >2.

Folgende Aspekte der Her-2/neu-Bestimmung wurden in Absprache mit der Vertreterin der Expertengruppe für histopathologische Befundung einer Evidenzrecherche und –bewertung unterzogen:

1. Ansprechen auf die Therapie mit Trastuzumab: Vergleich der Überexpression des Zellmembranrezeptors (IHC) und der Genamplifikation (FISH)
2. Reliabilität der Bestimmungsmethoden IHC und FISH
3. Konkordanz der Bestimmungsmethoden IHC und FISH
4. Reliabilität und diagnostische Sicherheit der CISH im Vergleich zu FISH
5. Her-2/neu-Status am definitiven Operationspräparat im Vergleich zur Bestimmung am Biopsiegewebe.

Diese Themen werden im Folgenden erläutert:

1. **Ansprechen auf die Therapie mit Trastuzumab: Vergleich der Überexpression des Zellmembranrezeptors (IHC) und der Genamplifikation (FISH)**

Die klinische Bedeutung des durch die jeweilige Bestimmungsmethode nachgewiesenen Her-2/neu-Status kann in Bezug auf drei Parameter gemessen werden:

1. die Signifikanz des Her-2/neu-Status als Prognosefaktor,
2. die diagnostische Sicherheit des Her-2/neu-Status in Bezug auf das Ansprechen auf Systemtherapien,
3. die diagnostische Sicherheit des Her-2/neu-Status in Bezug auf das Ansprechen auf die Therapie mit dem spezifischen Rezeptorantikörper Trastuzumab.

Ein Referenztest sollte alle drei Parameter am besten abbilden.

Für die vorliegende Analyse wurde die diagnostische Sicherheit der immunhistochemischen versus der FISH-Bestimmung in Bezug auf das Ansprechen auf die Trastuzumabtherapie einer Evidenzbewertung unterzogen.

2. **Reliabilität der Bestimmungsmethoden IHC und FISH**

In einer randomisierten kontrollierten Studie zum Beleg der Wirksamkeit einer Trastuzumabtherapie war 2000 eine große Variabilität der Her-2/neu-Rezeptor Bestimmung in verschiedenen Laboren aufgefallen [11]. Nur 88 von 119 Patientinnen waren auch bei zentraler Testung Her-2/neu positiv (IHC 3+ oder FISH+). Der lokale immunhistochemische Status wurde in 75% der Fälle bestätigt, der lokale FISH-Status in 67% der Fälle.

Vor diesem Hintergrund sollte die Frage nach der Reliabilität sowohl des immunhistochemischen als auch des FISH-Verfahrens einer Evidenzrecherche unterzogen werden.

3. **Konkordanz der Bestimmungsmethoden IHC und FISH**

Der vorliegende Algorithmus zur Her-2/neu-Bestimmung in Deutschland sieht eine FISH-Testung nur bei Vorliegen der immunhistochemischen Kategorie 2+ vor. Das heißt, für diese Kategorie dient die FISH als Referenzstandard. Es wird davon ausgegangen, dass die FISH-Bewertung für die anderen Kategorien (0/1+ und 3+) deshalb nicht erforderlich ist, da sich für diese Kategorien immunhistochemische Bewertung und Beurteilung durch FISH nicht unterscheiden.

Diese Annahme sollte einer systematischen Evidenzüberprüfung unterzogen werden, um die Änderung oder Beibehaltung des Algorithmus begründen zu können.

4. Reliabilität und Konkordanz der Methode CISH im Vergleich zu FISH

Ein Grund für den Einsatz von FISH in wenigen Laboren sind die Kosten und das aufwändige Verfahren. Die CISH-Methode weist ebenfalls eine Gen-Amplifizierung nach und ist einfacher und billiger durchzuführen. Dieses Verfahren wird in Finnland routinemäßig eingesetzt (pers. Kommunikation Y.Isola, 03/07). In den USA ist das Verfahren bisher nicht zugelassen [12]. In Deutschland wird die CISH-Methode bisher nur wenig angewandt.

Zur Evaluierung der diagnostischen Sicherheit der Methode sollte ein Vergleich mit der FISH-Methode erfolgen.

5. Her-2/neu-Bestimmung am definitiven Operationspräparat im Vergleich zur Bestimmung am Biopsiegewebe

Da die Gabe von Trastuzumab in Kombination mit einer Chemotherapie erfolgt und es zunehmend Protokolle für präoperativ verabreichte Chemotherapien gibt, gilt für die Her-2/neu-Bestimmung, dass sie für die präoperative Therapie aus histologischem Material von Biopsien vorgenommen werden muss.

In Bezug auf die unterschiedlichen Bestimmungsmethoden sollte geprüft werden, ob die Bestimmung aus Biopsiematerial zu den gleichen Ergebnissen führt, wie die Bestimmung am definitiven Operationspräparat.

Literatur

1. Carpenter G. The EGF receptor: a nexus for trafficking and signaling. *Bioessays*. 2000 Aug; 22(8):697-707 Sartore-Bianchi A, Ricotta R, Cerea G, Maugeri MR, Siena S. Rationale and clinical results of multi-target treatments in oncology. *Int J Biol Markers* 2007;22(1 Suppl 4):S77-87 Review
2. Gross ME, Shazer RL, Agus DB. Targeting the HER-kinase axis in cancer. *Semin Oncol* 2004;31(1 Suppl 3):9-20
3. Hoffmann F. Herceptin erhält EU-Zulassung für HER2-positiven Brustkrebs im Frühstadium. 2006. [cited: 2007 Aug 23]. Available from: <http://www.roche.com/de/med-cor-2006-05-24>
4. Baselga J, Perez EA, Pienkowski T, Bell R. Adjuvant trastuzumab: a milestone in the treatment of HER-2-positive early breast cancer. *Oncologist* 2006;11 Suppl 1:4-12. Review.
5. Goldhirsch A, Glick JH, Gelber RD, Coates AS, Thurlimann B, Senn HJ; Panel members. Meeting highlights: international expert consensus on the primary therapy of early breast cancer 2005. *Ann Oncol* 2005;16(10):1569-83
6. Pritchard KI, Shepherd LE, O'Malley FP, Andrulis IL, Tu D, Bramwell VH, Levine MN; National Cancer Institute of Canada Clinical Trials Group. HER2 and responsiveness of breast cancer to adjuvant chemotherapy. *N Engl J Med* 2006;354(20):2103-11
7. Deutsche Krebsgesellschaft (DKG). Interdisziplinäre S3-Leitlinie für die Diagnostik, Therapie und Nachsorge des Mammakarzinoms der Frau. München: Zuckschwerdt; 2004
8. Tanner M, Gancberg D, Di LA, Larsimont D, Rouas G, Piccart MJ, Isola J. Chromogenic in situ hybridization: a practical alternative for fluorescence in situ hybridization to detect HER-2/neu oncogene amplification in archival breast cancer samples. *Am J Pathol* 2000;157(5):1467-72
9. DAKO-Scoring-System; HercepTest
10. FISH-Scoring – System; Pathvysion-Test
11. Roche PC, Suman VJ, Jenkins RB, Davidson NE, Martino S, Kaufman PA, Addo FK, Murphy B, Ingle JN, Perez EA. Concordance between local and central laboratory HER2 testing in the breast intergroup trial N9831. *J Natl Cancer Inst* 2002;94(11):855-7
12. Carlson RW, Moench SJ, Hammond ME, Perez EA, Burstein HJ, Allred DC, Vogel CL, Goldstein LJ, Somlo G, Gradishar WJ, Hudis CA, Jahanzeb M, Stark A, Wolff AC, Press MF, Winer EP, Paik S, Ljung BM; NCCN HER2 Testing in Breast Cancer Task Force. HER2 testing in breast cancer: NCCN Task Force report and recommendations. *J Natl Compr Canc Netw* 2006;4 Suppl 3:S1-22; Review

2. Fragestellung

Folgende Schlüsselfragen wurden von der Methodikerin in Absprache mit der Vertreterin der Expertengruppe für pathohistologische Befundung generiert:

1. Sprechen Patientinnen mit Her2-Rezeptor – Überexpression auf der Zelloberfläche in gleichem Maße auf eine Therapie mit Trastuzumab an, wie Patientinnen mit einer Amplifikation des entsprechenden Gen-Abschnitts auf Chromosom 17?
2. Inwieweit sind die Methoden IHC und FISH reliabel?
3. Ist die Methode IHC und FISH im Hinblick auf ein negatives (0,1+) oder positives (3+) Ergebnis konkordant mit der FISH?
4. Inwieweit ist die Methode CISH reliabel und konkordant mit der FISH-Bestimmung?
5. Entspricht die diagnostische Wertigkeit der Her2- Bestimmung durch IHC / FISH / CISH aus Biopsiematerial der einer Bestimmung aus dem definitiven Operationspräparat?

3. Methoden

3.1 Kriterien für den Einschluss- und Ausschluss von Studien

Die folgenden Kriterien wurden vorab festgelegt.

3.1.1 Zielgruppe

Zielgruppe waren Patientinnen mit invasivem Mammakarzinom. Für Patientinnen mit duktalem in-situ Karzinom ist die Bedeutung der Her-2/neu-Expression noch nicht geklärt [5].

3.1.2 Diagnostische Verfahren

Die untersuchten diagnostischen Verfahren waren die Her-2/neu-Bestimmung mit den Methoden Immunhistochemie, Fluoreszenz-in-situ-Hybridisierung und Chromogene-in-situ-Hybridisierung. Als Referenzstandard wurde die Fluoreszenz-in-situ-Hybridisierung festgelegt.

Für Frage 2 und 3 wurde darüber hinaus Folgendes festgelegt:

Analysiert wurden Studien mit Verwendung folgender immunhistochemischer Testkits oder Antikörper:

- Testkits: Herceptest (DAKO)
- die kommerziellen monoklonalen Antikörper : CB11, TAB 250
- der kommerziellen polyklonalen Antikörper: A0485 (DAKO) sowie der
- polyklonale Antikörper R60 und
- der Test aus den Herceptin-Zulassungsstudien: CTA (Clinical Trial Assay = Kombination aus 4D5 und CB11)

3.1.3 Studientypen

Für die vorliegenden Fragestellungen des Vergleichs von Reliabilität und Konkordanz von Methoden sind prospektive, konsekutive multizentrische Kohortenstudien mit verblindeter Durchführung der jeweiligen Test-Methoden am besten geeignet. In die Bewertung wurden auch retrospektive und nicht konsekutive Studien sowie nicht verblindete Studien aufgenommen. Der Evidenzgrad für diese Studien wurde entsprechend geringer bewertet (siehe 3.4). Einzelfallberichte wurden ausgeschlossen.

3.1.4 Erforderliche Angaben

- Angabe der verwendeten Reagenzien (Hersteller)
- Für das jeweilige Verfahren müssen die absoluten Zahlen (positiv/negativ) in der Studie so angegeben sein, dass ein Vergleich der Verfahren möglich ist
- Angabe des verwendeten Materials

Die Fixierung des Gewebes war bei vielen Studien nicht angegeben, ebenso fehlten häufig Angaben über die Qualitätssicherungsmaßnahmen im Labor. Diese Angaben wurden deshalb in die Tabellen aufgenommen, aber nicht als zwingend vorausgesetzt. Das Fehlen dieser Angaben beeinträchtigt jedoch die Validität der Vergleiche der Studien.

3.1.5 Einschluss- und Ausschlusskriterien

Es galten folgende **Einschlusskriterien**:

1. Studie/Review beantwortet die Fragestellung
2. adäquate Studienpopulation
3. erforderliche Angaben sind genannt

Alle Einschlusskriterien mussten erfüllt sein.

Es galten folgende **Ausschlusskriterien**:

1. Studie wertet zytologische Befunde aus (Feinnadelpunktion)
2. Studie untersucht Her-2/neu-Bestimmungen aus dem Serum, mittels ELISA oder PCR
3. Verwendung nicht gebräuchlichen Materials zur Her-2/neu-Bestimmung
4. Methodisch: Einzelfallberichte
5. Doppelpublikationen
6. Für Frage 3: Studie wertet weniger als 50 Fälle aus

Kein Ausschlusskriterium durfte vorliegen.

3.2. Literaturrecherche

Die Literaturrecherche erfolgte nach Absprache mit dem Auftraggeber ohne zeitliche Beschränkung. Rechercheschlussdatum war der 15.2.2007 für Frage 2 und 3 sowie der 16.3.2007 für Frage 1, 4 und 5. Eine Publikation vom 7.5.2007 wurde für Frage 2 nachträglich aufgenommen.

Die Abstracts der Literatursuche wurden von der Methodikerin auf inhaltliche Relevanz geprüft, nicht relevante Abstracts wurden ausgeschlossen. Inhaltlich als relevant bewertete Abstracts wurden als Volltexte beurteilt. Der Ausschluss von Volltexten ist jeweils begründet.

Nach Vorgabe des Auftraggebers wurde im ÄZQ eine Literaturrecherche in folgenden Datenbanken durchgeführt:

1. PubMed (Internetportal der National Library of Medicine)
(<http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed>)
2. Database of Abstracts of Reviews of Effects (DARE), des Centre for Reviews and Dissemination (CRD)
(<http://www.york.ac.uk/inst/crd/crddatabases.htm#DARE>)
3. Cochrane-Datenbanken 'Cochrane Reviews' und 'Clinical Trials'
(<http://www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/HOME>).

Folgende Suchbegriffe wurden verwendet:

a) Für PubMed:

zu Frage 1) (clinical outcome OR response OR time to progression) AND (HER2 OR HER 2 OR c-erbB2) AND (FISH OR fluorescen* OR CISH OR chromogen*) AND breast AND (IHC OR immunohistochemi*)

zu Frage 2 und 3) (Her-2/neu OR HER 2 OR Her2 OR c-erbB2) AND 'breast' AND (FISH OR fluoresc*) AND (IHC OR immunohistochemi*)

und (Her-2/neu OR HER 2 OR Her2 OR c-erbB2) AND breast AND (FISH OR fluoresc*) AND (IHC OR immunohistochemi*) AND (concordan* OR perform* OR accura* OR reliab*)

zu Frage 4) 'breast' AND (CISH OR chromogen*) AND (FISH OR fluoresc*)

zu Frage 5) biops* AND (core OR vacuum*) AND breast AND (HER2 OR Her 2 OR c-ErbB2) AND (accura* OR reliab* OR concord*)

b) Für DARE und Cochrane

zu Frage 1-3) breast AND fluorescen* (Cochrane); breast AND HER2 (DARE)

zu Frage 4) breast AND CISH

zu Frage 5) breast AND biops* AND Her 2

Die Suchbegriffe wurden jeweils als Textwort verwendet, d.h. nicht nur auf Titel oder Abstract bezogen. Die Verwendung des MeSh-Schlagwortes ‚breast neoplasms‘ führte zu erheblich weniger (aktuellen) Zitaten, da die Verschlagwortung der Artikel einige Monate in Anspruch nehmen kann. Auf die Verwendung des MeSh-Schlagwortes wurde deshalb verzichtet. Ebenso wurde auf ‚Cancer‘ verzichtet, da die Suchergebnisse dadurch nur um jeweils weniger als fünf Zitate gemindert wurden.

3.3 Zielgrößen und deren Definition

Folgende Zielgrößen wurden festgelegt:

Zu Frage 1

Hier galten als Zielgrößen die Kriterien für das Ansprechen auf eine Therapie bei metastasiertem Brustkrebs, da Studien zum Therapieansprechen nur für Patientinnen mit metastasiertem Brustkrebs vorliegen.

Zu Frage 2 und 4

Hier wurde die Zuverlässigkeit der diagnostischen Verfahren IHC und FISH analysiert. Als Zielgröße galt deshalb die Realibilität der Verfahren, ermittelt aus dem Anteil der Variabilität der Befundungen [13], entweder bezogen auf mehrere Befunder in einem Labor oder bezogen auf Befunder aus verschiedenen Laboren.

⇒ Die Reliabilität wurde wie folgt definiert

$$\frac{\text{Anzahl als gleich befundene Gewebeprobe durch unterschiedliche Befunder}}{\text{Gesamtzahl der Beurteilungen}}$$

Diese Angaben beziehen sich jeweils auf eine Methode.

Neben dieser prozentualen Übereinstimmung kann die Übereinstimmung von Befunden auch mit dem Kappa-Wert ausgedrückt werden. Der Kappa-Wert gibt an, inwieweit die Übereinstimmung von Befunden ein rein zufallsbedingtes Maß übersteigt. Der Kappa-Wert wurde aufgeführt, falls er in der Studie angegeben war.

Er wurde nach Landis und Koch bewertet [14]. Es gelten folgende Kategorien: $\kappa < 0$ = „schlechte Übereinstimmung (poor agreement)“, κ zwischen 0 und 0,20 = „etwas (slight) Übereinstimmung“, 0,21-0,40 = „ausreichende (fair) Übereinstimmung“, 0,41-0,60 = „mittelmäßige (moderate) Übereinstimmung“, 0,61-0,80 = „beachtliche (substantial) Übereinstimmung“, 0,81-1,00 = „(fast) vollkommene (almost perfect) Übereinstimmung“.

Zu Frage 3

Die spezifische Therapie bei positivem Her-2/neu-Status kann bei nachgewiesener Rezeptorüberexpression oder bei Amplifikation des Genabschnitts auf Chromosom 17 erfolgen. Es interessierte deshalb die Konkordanz der beiden gängigen Methoden IHC und FISH.

⇒ Die Konkordanz für die beiden Methoden wurde wie folgt definiert:

$$\frac{\text{Anteil der durch unterschiedliche Methoden als gleich beurteilten Gewebeprobe}}{\text{Gesamtzahl der Beurteilungen}}$$

Zu Frage 4

In Bezug auf das CISH-Verfahren wurde primär die Konkordanz zu FISH als der Referenzmethode der in situ Hybridisierung bestimmt. Der Kappa-Wert wurde, ebenso wie die Sensitivität und Spezifität der Methode in Bezug auf FISH aufgeführt, falls in den Studien dazu Angaben waren. Darüber hinaus war die Reliabilität der Methode Zielgröße.

Zu Frage 5:

Hier wurde die Konkordanz der Her-2/neu-Bestimmung am Biopisematerial im Vergleich zu der Bestimmung am definitiven Operationspräparat geprüft.

3.4 Bewertung der gefundenen Literatur

Für die Bewertung der vorliegenden Studien wurde die Oxford-Klassifikation für Diagnosestudien verwendet. Für die Evidenzstufe 2 bedarf es einer verblindeten Anwendung der Tests sowie einer konsekutiven Reihe an untersuchten Fällen, fehlt eines dieser Kriterien, wird die Evidenzstufe 3 vergeben. Die Studien wurden grundsätzlich als Kohortenstudien gewertet.

Literatur

13. Gordis L. Epidemiologie. Marburg: Kilian-Verlag; 2001
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-174

4. Ergebnisse

4.1 Vergleich der HER-2/neu-Bestimmung durch IHC und FISH in Bezug auf das Ansprechen auf die Therapie mit Trastuzumab

4.1.1 Ergebnisse der Literaturrecherche

Die Recherche in DARE erbrachte 23 Zitate, die Suche in Cochrane ergab 31 Zitate. Es wurden drei Dubletten zu PubMed identifiziert, die restlichen Abstracts waren inhaltlich nicht relevant.

Rechercheergebnisse für PubMed

PubMed bis 16.3.2007	n= 156
Nach Screenen der Abstracts bewertete Volltexte	n= 14
Reviews	n= 6
Systematische Reviews	n= 0
Einzelstudien zum Thema	n= 7
Handsuche in Literaturverzeichnissen	n= 1
Eingeschlossene Einzelstudien	n= 4

4.1.2 Aufbereitete Evidenz

Es wurde kein systematischer Review zu der Fragestellung identifiziert.

4.1.3 Evidenz von Einzelstudien

a) Studien- und Publikationsqualität

Die Recherche zu dieser Fragestellung erbrachte sieben Einzelstudien. Ausgeschlossen wurden drei Studien (siehe 5.1.1 und 5.1.2.).

Bei einer der eingeschlossenen Studien wurde der Her-2/neu-Status prospektiv durch FISH und immunhistochemisch durch vier verschiedene Antikörper bestimmt (Seidman et al., 2001). In dieser Studie war die statistische Signifikanz für das Therapieansprechen pro Bestimmungsmethode angegeben.

Bei zwei weiteren Studien (Vogel et al., 2002; Mass et al., 2005) wurde der Her-2/neu-Status prospektiv mit IHC bestimmt und die FISH-Bestimmung erfolgte zu einem späteren Zeitpunkt aus Archivmaterial (beide verwendeten zur immunhistochemischen Bestimmung Antikörper, die kommerziell nicht erhältlich sind). Hofmann et al., 2007 wendeten sowohl die IHC als auch die FISH auf Archivmaterial an.

In diesen drei Studien wurde der prozentuale Anteil der Fälle, die auf die Therapie ansprachen, für eine oder beide Bestimmungsmethoden angegeben.

Die Studien wurden mit einem Evidenzgrad von 3b bewertet, für die prospektive Studie von Seidman et al. mit durchgehend verblindeter Beurteilung wurde ein Evidenzgrad 2b vergeben.

Hinsichtlich der Publikationsqualität ist anzumerken, dass in drei der Studien die Art der Fixierung der Gewebeproben nicht angegeben war.

b) Inhaltliche Ergebnisse

Die vorliegenden Studien (n= 95 bis n= 784) prüften die Wirksamkeit der Trastuzumab-Therapie als Monotherapie (Vogel et al., 2002; Hofmann et al., 2007, Mass et al., 2005 Studie 2) oder in Kombination mit einer Chemotherapie (Seidman et al., 2001; Mass et al., 2005, Studie1) bei metastasiertem Brustkrebs.

Die prospektive Studie von Seidman zeigte, dass sowohl der positive Her-2/neu-Status aufgrund der IHC als auch der positive FISH-Status statistisch signifikant für das Ansprechen auf Trastuzumab kombiniert mit

Paclitaxel war. Als Her-2/neu positiv wurden hierbei alle als IHC2+ und IHC3+ befundenen Karzinomgewebe gewertet. Es wurde nicht nach IHC 2+ oder IHC3+ stratifiziert und es ist nicht bekannt, welchen Anteil IHC2+ in Bezug auf alle, als positiv bewerteten Fälle, hatte. Beim Vergleich der immunhistochemisch verwendeten Antikörper zeigte der positive Herceptest die geringste statistische Signifikanz ($p=0,03$) in Bezug auf das Therapieansprechen. Bei einer im Vergleich zu den anderen eingesetzten Antikörpern hohen Rate an Her-2/neu positiven Patientinnen muss hier von einem erhöhten Anteil falsch positiver Befunde ausgegangen werden. Die höchste Signifikanz wurde bei Einsatz des Antikörpers TAB250 erzielt ($p=0,004$). Dieser Wert wurde auch für die FISH erreicht.

Die Studie von Vogel et al. wies ein gering höheres Ansprechen bei IHC3+ (35%) als bei FISH+ (34%) für die Trastuzumab-Monotherapie auf.

In dieser Studie wurde bei einer FISH negativen, aber IHC positiven (3+) Patientin ein Ansprechen auf Trastuzumab gezeigt, dies entsprach einer Ansprechrate von 7% bei FISH-negativen Patientinnen. Für immunhistochemische Befunde der Kategorie IHC2+ konnte kein Therapieansprechen gezeigt werden.

Hofmann et al. fanden, dass alle Patientinnen, die auf die Trastuzumabtherapie angesprochen hatten, IHC 3+ positiv waren, aber nur 84% waren FISH positiv. Das heißt, drei Patientinnen, die IHC3+ und FISH negativ waren, sprachen auf die Trastuzumabtherapie an. Bei zwei dieser Patientinnen lag eine Polysomie des Chromosoms 17 vor.

Mass et al. zeigten in einer nachträglich durchgeführten FISH-Analyse ein signifikant besseres Therapieansprechen auf Trastuzumab für FISH positive Patientinnen (78%) im Gegensatz zu FISH negativen Patientinnen (22%). Darüber hinaus zeigte sich bei FISH-positiven Patientinnen eine signifikant längere Zeit bis zum Auftreten einer Krankheitsprogression und ein signifikant verbessertes Gesamtüberleben. In diese Studie waren Patientinnen mit IHC2+ und IHC3+ eingeschlossen worden. Der Anteil an IHC2+ bei den FISH-negativen Patientinnen lag bei 72%. Die Ansprechrate der Patientinnen mit ausschließlich IHC3+ ist in der Studie nicht angegeben.

4.1.4 Diskussion

a) Methodische Diskussion

Zu der Frage des Therapieansprechens in Bezug auf die Ergebnisse der Bestimmungsmethoden FISH und IHC liegen nur Daten aus wenigen – und mit einer Ausnahme – retrospektiven Studien zu metastasiertem Brustkrebs vor, nicht jedoch für die adjuvante Situation.

Methodisch sind die Studien im Hinblick auf die Aussage zur IHC- bzw. FISH-Bestimmung durch unterschiedliche Vorbehandlungen und verabreichte Therapieregime nicht ohne weiteres vergleichbar. Auch die Art der Gewebefixierung kann die Ergebnisse beeinflussen, diese ist bei drei der vier Studien nicht angegeben. Für eine Validierung der Ergebnisse sind deshalb prospektive Studien erforderlich.

b) Inhaltliche Diskussion

Für die Beantwortung der vorliegenden Frage ergab die einzige vorliegende prospektive Studie für beide Bestimmungsverfahren ein statistisch signifikantes Ergebnis für das Therapieansprechen bei positivem Her-2/neu-Status.

Die größte verfügbare Studie mit einer Studienpopulation von Her-2/neu-positiven Patientinnen aufgrund von Immunhistochemie 2+ und 3+ zeigte demgegenüber eine sehr viel bessere diagnostische Sicherheit für die nachträglich durchgeführte FISH-Analyse. Eine mögliche Erklärung für die bessere Diskriminierung von FISH+ in Bezug auf das Therapieansprechen könnte darin liegen, dass die FISH negativen Patientinnen zu mehr als zwei Dritteln als IHC2+ befundet worden waren. Für die Ergebnisse dieser immunhistochemischen Kategorie wurde in einer weiteren Studie gezeigt, dass sie nicht auf die Trastuzumabtherapie ansprechen, wenn sie nicht FISH positiv sind.

Für das Ansprechen von IHC3+, die FISH negativ sind, liegen Daten von vier Patientinnen vor, die zeigen, dass bei einer starken Rezeptorüberexpression ein Therapieansprechen auch ohne Genamplifikation möglich ist. Bei zwei der vier Fälle lag eine Polysomie des Chromosoms 17 vor.

Insgesamt kann aus der vorliegenden Datenlage nur eine eingeschränkt valide Aussage zum Therapieansprechen getroffen werden. Die derzeitige Vorgehensweise, sowohl IHC3+ als auch FISH positive Patientinnen für eine Therapie mit Trastuzumab zuzulassen, wird durch die Studien in der Tendenz bestätigt.

Um ein sicheres Testprocedere festzulegen sind weitere Studien mit höherer statistischer Power im Hinblick auf das Ansprechen von Trastuzumab erforderlich, insbesondere mit den jetzt gebräuchlichen Antikörpern und Testkits.

4.2 Reliabilität der Methoden IHC und FISH

4.2.1 Ergebnisse der Literaturrecherche

Die Suche in DARE und Cochrane (gleiche Anzahl Zitate wie unter 4.1.1, da gleiche Suchstrategie) ergab für diese Fragestellung eine Dublette zu PubMed und keine weiteren inhaltlich relevanten Zitate.

Rechercheergebnisse für PubMed

PubMed bis 27.2.2007	n= 751
Nach Screenen der Abstracts geprüfte Volltexte:	n= 21
Reviews	n= 0
Einzelstudien	n= 21
Davon eingeschlossene Studien	n= 19

4.2.2 Aufbereitete Evidenz

Es wurde kein Review zu der Fragestellung identifiziert.

4.2.3 Evidenz von Einzelstudien

a) Studien- und Publikationsqualität

Zu der Frage der Reliabilität von IHC und FISH ergab die Recherche insgesamt 21 Studien. Davon erfüllten 19 die Einschlusskriterien, acht davon waren prospektiv.

In Bezug auf die Publikationsqualität ist festzustellen, dass in sieben Studien nicht angegeben wurde, ob es sich um verblindete Vergleiche handelte.

Ebenso war die Angabe, ob das jeweilige Ergebnis anhand von mitgelaufenen Kontrollen überprüft wurde, nur in acht Studien angegeben. Die Studien wurden mit 2b bewertet, wenn sie verblindet durchgeführt wurden und eine Ergebnisüberprüfung durch entsprechende Kontrollen erfolgte, ansonsten mit 3b. Da nur wenige Studien einen Kappa-Wert angaben, wird dieser im Folgenden nicht dargestellt.

b) Inhaltliche Ergebnisse

Reliabilität für die FISH

Befunder im gleichen Labor

Die Reliabilität in Bezug auf mehrere Befunder im gleichen Labor kann im Median mit 96,6% (95%-98%) angegeben werden (Bartlett et al.; Press et al.; Tsuda et al.; 2001).

Laborvergleich

Die Reliabilität im Vergleich von lokalen zu zentralen Laboren beträgt im Median 88,1% (67%-92%) (Press et al.; Perez et al.; Roche et al.).

Für die Übereinstimmung zwischen zentralen Laboren wurden 95,2% und 98,7% berichtet (Press et al.; Perez et al.). Dowsett et al. erreichten in einem internationalen Ringversuch zwischen fünf Referenzlaboren 80% absolute Übereinstimmung der Befunde.

Die meisten diskordanten Befunde wurden für eine niedrige Genamplifikation (low level Bereich zwischen 1.7-2.3 Genkopien) gefunden.

Reliabilität für die IHC

Befunder im gleichen Labor

Die Reliabilität des immunhistochemischen Verfahrens im Hinblick auf mehrere Befunder im gleichen Labor betrug stratifiziert nach 0 bis 3+ im Median 84% (48%-96%) (Bartlett et al.; Hsu et al.; Tsuda et al. 2002; Yaziji et al.).

Wird nur nach der Bewertung positiv oder negativ stratifiziert, liegen die Werte zwischen 62% und 96%.

Bei Vergleich der am häufigsten eingesetzten Antikörper Herceptest und Cb11 zeigten sich in Bezug auf die Reliabilität sowohl ein besseres als auch ein schlechteres Ergebnis für den Herceptest (Bartlett et al.; Tsuda et al. 2002).

Laborvergleich

Für die Übereinstimmung der Befunde zwischen lokalen und zentralem Labor ergab sich im Median 75% (40%-81,6%) (Press et al.; Paik et al.; Perez et al.; Reddy et al.; Rhodes et al.).

Für zentrale Labore fand sich in dem internationalen Ringversuch mit fünf Laboren von Dowsett et al. eine absolute Übereinstimmung von 45% und in einer retrospektiven Auswertung mit drei Laboren 66%.

Perez et al. berichteten von einer Übereinstimmung von 94,3% zwischen zentralem und Referenzlabor.

In Bezug auf die einzelnen immunhistochemischen Tests wurden ebenfalls heterogene Ergebnisse erzielt. Besonders schlecht reproduzierbar waren IHC2+ Ergebnisse (Bilous et al., 2002; Reddy et al.).

4.2.4 Diskussion

a) Methodische Diskussion

Für die Fragestellung der Reliabilität von IHC und FISH wurden retrospektive und prospektive Studien identifiziert. Eine verblindete Beurteilung war nicht durchgängig angegeben, ebensowenig die Maßnahmen zur Überprüfung der Ergebnisse. In der Tendenz zeigen aber Studien höherer und minderer Qualität ein einheitliches Gesamtergebnis in Bezug auf die Unterschiede zwischen den beiden Verfahren.

b) Inhaltliche Diskussion

Die FISH zeigte sowohl hinsichtlich der Übereinstimmung zwischen mehreren Befundern als auch in Bezug auf die Übereinstimmung zwischen mehreren Laboren konsistent bessere Ergebnisse.

Die immunhistochemische Bestimmung wies eine sehr große Variabilität auf. Auch für einzelne Antikörper war das Ergebnis heterogen. Die Überlegenheit eines bestimmten Antikörpers oder Testkits kann aus den Studien nicht abgeleitet werden. Aus einer hohen Übereinstimmung von Daten eines zentralen und eines Referenzlabors ergibt sich ein Hinweis auf eine mögliche Verbesserung der Reliabilität von IHC unter standardisierten Bedingungen. Weitere Faktoren für die Verbesserung der Reliabilität beider Verfahren wie zum Beispiel die einheitliche Gewebeaufbereitung oder die eingesetzten Kontrollen, sind zu prüfen.

4.3 Konkordanz von IHC und FISH

4.3.1 Ergebnisse der Literaturrecherche

Die Recherche in DARE und Cochrane (gleiche Anzahl Zitate wie unter 4.1.1, da gleiche Suchstrategie siehe 4.1) ergab neben drei nicht verfügbaren Publikationen sieben Dubletten zu PubMed.

Rechercheergebnisse für PubMed

PubMed bis 16.3.2007	n= 751
Nach Screenen der Abstracts geprüfte Volltexte	n= 79
Systematische Reviews	n= 1 (nach Rechercheabschluss aufgenommen)
Einzelstudien zum Thema	n 79
Eingeschlossene Einzelstudien	n= 47 in 2 Tabellen
Ausgeschlossene Studien	n= 32

4.3.2. Aufbereitete Evidenz

Es wurde nachträglich ein systematischer Review zur Konkordanz IHC-FISH eingeschlossen: Dendukuri et al., 2007: Testing for Her2-positive breast cancer: a systematic review and cost-effectiveness analysis. Dieser Review wurde Anfang Mai 2007 veröffentlicht, als die Recherchen schon abgeschlossen waren.

Die Ergebnisse des Reviews werden im Folgenden dargestellt.

Es erfolgte eine systematische Literaturrecherche bis Juni 2006 mit transparenten Einschlusskriterien. Eingeschlossen wurden 17 Studien, Voraussetzung war die Verwendung eines in Kanada lizenzierten IHC-Tests. Die Gewebeproben mussten konsekutiv oder per Zufall ausgewählt worden sein.

Es wurde die Frage beantwortet, wie der Anteil positiver und negativer immunhistochemischer Kategorien bei den Mammakarzinomen verteilt ist, und welche Konkordanz sich bei den immunhistochemischen Ergebnissen zu denen von FISH zeigt. Des Weiteren wurde eine Kosten-Effektivitätsanalyse in Bezug auf den Einsatz von FISH ergänzend zu IHC durchgeführt.

⇒ Folgende Ergebnisse wurden ermittelt:

1. Die mediane Verteilung von IHC-Scores in Brustkrebsgewebe:

0	=	36,1%
1+	=	35,5%
2+	=	12,0%
3+	=	16,2%

2. Der mediane Anteil FISH-positiver Fälle pro IHC-Score:

0	=	1,6%
1+	=	4,9%
2+	=	29,8%
3+	=	92,4%

Bei FISH als Goldstandard ergab sich als kostengünstigste Strategie, IHC2+ und IHC3+ Fälle zusätzlich durch FISH zu testen.

4.3.3 Evidenz von Einzelstudien

a) Studien- und Publikationsqualität

Die vorliegende Fragestellung der Konkordanz von IHC und FISH wurde in zahlreichen Studien untersucht. Für die vorliegende Auswertung wurden nur Studien mit 50 oder mehr Fällen berücksichtigt, um pro immunhistochemischer Kategorie – insbesondere für die Kategorie 3+ – nicht nur Einzelfälle zu bewerten, da die IHC-Befundung eine hohe Variabilität aufweist.

Für die vorliegende Fragestellung der Konkordanz zu FISH und IHC 3+ konnten nur Studien verwendet werden, die zwischen den IHC-Kategorien 2+ und 3+ bei als ‚positiv‘ angegebenen Werten stratifizierten. Für die Konkordanz 0/1+ wurden die beiden Werte als ‚negativ‘ zusammengefaßt.

Konsekutive Studien mit verblindeter Befundung wurden mit 2b bewertet, fehlte eines der Merkmale, wurde ein Evidenzgrad von 3b vergeben.

b) Inhaltliche Ergebnisse

Die Konkordanz IHC3+ zu FISH+ wurde für die vorliegenden Einzelstudien zunächst im Hinblick auf die zur immunhistochemischen Bestimmung verwendeten Antikörper bzw. das verwendete Testkit stratifiziert.

Herceptest

Das am häufigsten zum Einsatz kommende Testkit war der Herceptest (Hersteller: DAKO). Dieser wurde in 23 der vorliegenden Studien verwendet. Die Anzahl an IHC3+ in diesen Studien betrug zwischen n= 12 und n= 707.

Der mediane Anteil FISH positiver Bestimmungen bei IHC3+-Status für diese Studien lag bei 90,9% (38%-100%). Die berechnete gepoolte Konkordanz ergab einen Wert von 88,09% (1724/1957).

Für die Kategorie 0/1+ lag der mediane Anteil Übereinstimmung zu FISH – bei 97%.

CB11-Antikörper

Bei Verwendung des CB11-Antikörpers lag der mediane Anteil FISH positiver Proben in der Kategorie IHC3+ bei 85% (71,4%-100%) unter Anwendung des DAKO-Systems (vier Studien). Die Berechnung der gepoolten Konkordanz ergab 79,8% (75/95).

A0485-Antikörper

Für A0485 zeigte sich in drei Studien (85% und zwei Mal 100%, d.h. eine mediane Konkordanz von 100%) eine gepoolte Konkordanz von 96% (50/52).

Übereinstimmung der Ergebnisse unabhängig vom eingesetzten Test

Der mediane Anteil an Fish+ bei IHC 3+ lag insgesamt bei 90%. Die gepoolte Gesamtkonkordanz aus 29 Studien – unabhängig vom verwendeten AK – beträgt: 88,3% (2.373/2.686).

Der mediane Anteil an FISH negativen Proben bei IHC0/1+ lag bei 98,2% (91%-100%).

Die gepoolte Konkordanz von IHC 0/1+ zu FISH betrug bei den Studien (alle verwendeten Antikörper eingeschlossen) 96,1% (6.105/6.354).

In vier Studien wird nach 0 und 1+ stratifiziert. Für 0 ergibt sich eine gepoolte Konkordanz von 97,3% (2.295/2.359), für 1+ eine gepoolte Konkordanz von 94% (2.271/2.419).

4.3.4 Diskussion**a) Methodische Diskussion**

Für die Fragestellung nach der Konkordanz von IHC und FISH am gleichen Gewebe wurde nach Abschluss der Recherchen ein systematischer Review identifiziert.

Diesem wird in der Folge das Ergebnis aus den Einzelstudien gegenübergestellt. Die eingeschlossenen Einzelstudien konnten für die Beantwortung der Frage nach der Konkordanz bei IHC3+ zu FISH + nur zum Teil verwendet werden, da eine Stratifizierung nach 2+ und 3+ nicht durchgehend erfolgte. Insgesamt wurden zur Berechnung der Konkordanz IHC-FISH – unabhängig vom verwendeten Antikörper- 29 Studien einbezogen. Von diesen wurden 12 auch in dem Review von Dendukuri et al. analysiert. Eine konsekutive Gewebeprobenauswahl oder eine Selektion ‚per Zufall‘ war für die vorliegende Auswertung im Gegensatz zum Review von Dendukuri et al. kein Einschlusskriterium, dies kann den Unterschied in der Zahl insgesamt eingeschlossener Studien erklären. Für Dendukuri et al. war dieses Kriterium vor allem deshalb wichtig, da sie zusätzlich zur Konkordanz auch die Verteilung von IHC-Scores in Brustgewebe ermitteln wollten.

Der Einfluss der Angabe verblindeter Untersuchungen bleibt zu prüfen.

b) Inhaltliche Diskussion

Der aktuelle Review von Dendukuri et al. zeigt eine mediane Konkordanz von 92,9% für IHC3+ und FISH anhand von 17 Einzelstudien. Die vorliegende Auswertung ergibt anhand von 29 Einzelstudien eine mediane Konkordanz von 90% für diese IHC-Kategorie.

Negative immunhistochemische Befunde (0/1) zeigen in der vorliegenden Auswertung von Einzelstudien eine sehr hohe gepoolte Konkordanz von 96,1% und eine mediane Konkordanz von 98,2%.

Dendukuri et al. fanden eine mediane Übereinstimmung von 98,4% bei IHC 0 und FISH negativen Proben und von 95,1% bei IHC 1+, die Werte aus der gepoolten Auswertung ergaben hier 97,3% und 94%.

In der Gesamtwertung zeigten die Kategorien IHC0/1+ sowohl in dem Review von Dendukuri als auch in den eingeschlossenen Einzelstudien eine sehr hohe Konkordanz und in der Kategorie IHC3+ eine vertretbare Konkordanz.

Nach der aktuellen ASCO-Leitlinie zur Her-2/neu-Bestimmung [15] wird aktuell vorgeschlagen, die IHC-Kategorien zu ändern und anstelle von 10% Zellen mit starker Überexprimierung von 30% auszugehen.

Dies könnte die Diagnosesicherheit für die IHC und demzufolge auch die Konkordanz zur FISH erhöhen.

Literatur

15. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, Pegram MD, Perez EA, Press MF, Rhodes A, Sturgeon C, Taube SE, Tubbs R, Vance GH, van de Vijver M, Wheeler TM, Hayes DF. American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer. J Clin Oncol 2007 Jan 1;25(1):118-45

4.4 Konkordanz und Reliabilität von FISH und CISH

4.4.1 Ergebnisse der Literaturrecherche

Für diese Fragestellung wurde kein relevantes Zitat in DARE und Cochrane gefunden.

Rechercheergebnisse für PubMed

PubMed-Ergebnisse bis 27.2.2007	n= 64
Nach Screenen der Abstracts geprüfte Volltexte	n= 21
Reviews	n= 0
Eingeschlossene Einzelstudien zum Thema Vergleich FISH-CISH	n= 21
Ausgeschlossene Studien	n= 0

4.4.2 Aufbereitete Evidenz

Zu dieser Fragestellung wurde kein systematischer Review identifiziert.

4.4.3 Evidenz von Einzelstudien

a) Studien- und Publikationsqualität

Bei dieser Fragestellung wurden keine Studien ausgeschlossen.

Aufgrund der eindeutig zu bestimmenden diagnostischen Parameter ‚negativ‘ oder ‚positiv‘ wurden auch kleine Fallzahlen zugelassen.

Die vorliegenden 20 Studien wurden formal alle mit 3b bewertet. 14 von 20 Studien wurden verblindet durchgeführt. Die Ergebnisse weichen von denen ohne angegebene Verblindung nicht ab.

Bezüglich der Bewertungskategorien gibt es leicht unterschiedliche cut-offs, die jeweiligen Ergebnisse wurden übernommen.

b) Inhaltliche Ergebnisse

Konkordanz FISH-CISH

Die Studien zeigten durchgängig eine gute bis hervorragende Konkordanz mit Kappa-Werten von 0.81-1.0 (Angabe aus 14 Studien).

Der gemittelte Kappa-Wert aus den 14 Studien beträgt 0.90. Bei Auswertung der Studien mit einem $n > 100$ beträgt der Wert 0.87 (0,81-1,0). Nach Landis und Koch zeigt ein Kappa-Wert über 0.81 eine exzellente Konkordanz.

Die prozentuale Korrelation positiv/negativ beträgt fast durchgängig mehr als 90%. Lediglich in drei Studien wurden Werte von 83,7%, 83,9% und 87,9% gemessen. Bilous et al. fanden Fälle von negativer CISH v.a. bei Nachweis von niedriger Amplifizierung durch FISH. Bei den anderen zwei Studien (Gupta et al., Diaz et al.) trat die Diskordanz bei negativer FISH und aufgrund von Polysomie des Chromosoms 17 schwach positiver CISH auf.

Die gepoolte Konkordanz aus den sechs Studien, die IHC2+ in absoluten Zahlen angeben, betrug 93,5% für FISH und CISH in dieser Kategorie.

Insgesamt 43x wurde FISH+ und CISH negativ angegeben, und 52x FISH- und CISH+, davon 16x bei mit FISH festgestellter Polysomie.

In zwei Studien wurden Angaben zur Reliabilität der CISH in Bezug auf die Bestimmung in mehreren Laboren gemacht: Bilous et al. (5 Labore) berichteten von einer Übereinstimmung in 79,6% und Tanner et al. (2 Labore) in 95,9%. Die Raten für die Übereinstimmung der Befundung im gleichen Labor betragen in zwei Studien 91% und 97%, in einer weiteren Studie wird ein Kappa-Wert von 0.96 angegeben.

4.4.4 Diskussion

a) Methodische Diskussion

Die vorliegenden Studien zeigen unabhängig von der Fallzahl eine große Übereinstimmung der Ergebnisse auf. Dies kann als Hinweis auf dessen Robustheit gewertet werden. Die Studienergebnisse waren des Weiteren unabhängig von einer angegebenen Verblindung der Verfahren. Auch die Studien, in denen keine Verblindung angegeben ist, stützen somit das Gesamtergebnis.

b) Inhaltliche Diskussion

Die Fragestellung nach der Konkordanz von FISH zu CISH kann aufgrund der vorliegenden Daten positiv beantwortet werden. Die angegebenen Kappa-Werte sprechen einheitlich für eine exzellente Konkordanz zwischen den beiden Verfahren.

Für die Diskriminierung der Her-2/neu Genamplifikation in positiv und negativ können das CISH-Verfahren und das FISH-Verfahren demzufolge als gesichert konkordant gelten.

Die prozentuale Übereinstimmung der beiden Verfahren bei Bestimmung aus Geweben, die immunhistochemisch als 2+ bewertet wurden, spricht für eine gute Detektionsrate von CISH auch im niedrig amplifizierten Bereich, denn FISH-positive Fälle, die immunhistochemisch mit IHC2+ bewertet werden, sind häufig niedrig amplifiziert. Dennoch ist die Konkordanz in diesem Bereich am schlechtesten.

Die Reliabilität der CISH erscheint der Reliabilität der FISH vergleichbar.

4.5 Konkordanz der Her2/neu-Bestimmung aus Biopsiegewebe im Vergleich mit dem definitiven Operationspräparat

4.5.1 Ergebnisse der Literaturrecherche

Die Recherche in DARE und Cochrane erbrachte für diese Fragestellung keine relevante Publikation.

Rechercheergebnisse für PubMed

PubMed bis 16.3.2007	n= 19
Nach Screenen der Abstracts geprüfte Volltexte	n= 12
Reviews	n= 0
Einzelstudien zum Thema	n= 11; Handsuche: n= 1
Eingeschlossene Einzelstudien	n= 9

4.5.2 Aufbereitete Evidenz

Die Recherche erbrachte für diese Fragestellung keinen Review.

4.5.3 Evidenz von Einzelstudien

a) Studien- und Publikationsqualität

Zu der vorliegenden Fragestellung nach der Übereinstimmung der Ergebnisse des Her-2/neu-Status am Biopsiematerial und am definitiven Operationspräparat wurden 12 Studien identifiziert, davon wurden drei Studien ausgeschlossen (siehe 5.5.1. und 5.5.2).

Von den eingeschlossenen neun Studien wurden lediglich zwei (Cahill et al.; Jacobs et al.) prospektiv durchgeführt. Die Fallzahl lag bei sechs der Studien unter n= 100.

Sechs Studien gaben eine konsekutive Aufarbeitung von Gewebeproben an.

In zwei Studien war ein verblindeter Vergleich angegeben, für diese Studien wurde ein Evidenzgrad 2b gegeben.

In sechs Studien fehlte diese Angabe. Jacobs et al. arbeiteten nicht verblindet. Für diese Studien wurde ein Evidenzgrad von 3b vergeben.

b) Inhaltliche ErgebnisseKonkordanz Biopsie zu definitivem Operationspräparat für die IHC

Das immunhistochemische Ergebnis aus dem Biopsiematerial zeigte im Vergleich zum definitiven Operationspräparat sowohl vorwiegend eine höhere Kategorie (Taucher et al.; Müller-Holzner et al.; Mann et al.; Cahill et al.: insgesamt in 32 Fällen) als auch vorwiegend eine niedrigere Kategorie (Erggelet et al.; Burge et al.: insgesamt in 23 Fällen). Nur eine Studie zeigte eine vollständige Übereinstimmung.

Stratifiziert in Bezug auf die Einzelergebnisse ergab sich der höchste Anteil konkordanter Bestimmungen bei IHC 3+ (median 89%) während für die beiden anderen Kategorien (0/1+ und 2+) 77,8% und 78,8% ermittelt wurden.

Die Studie von Taucher et al. zeigte bei IHC3+ am Biopsiematerial und negativem immunhistochemischen Ergebnis am definitiven Operationspräparat in 85% der Fälle auch eine negative FISH.

Konkordanz Biopsie zu definitivem Operationspräparat für die FISH

Daten zur Konkordanz von FISH aus CNB und FSE liegen nur aus der Studie von Burge et al. aus n= 18 Bestimmungen vor und betragen 100%.

4.5.4 Diskussion**a) Methodische Diskussion**

Für die Beantwortung der Frage nach der Übereinstimmung der Her-2/neu-Ergebnisse in Bezug auf Biopsie und definitives Operationspräparat liegen nur drei größere Studien (n =/ >100) vor. Die Beurteilung der Ergebnisse wird durch die häufig fehlende Angabe der Verblindung eingeschränkt. Daten zur FISH-Bestimmung liegen nur aus einer Studie vor. Die Frage kann aufgrund der Studienlage nur eingeschränkt valide beantwortet werden.

b) Inhaltliche Diskussion

Die vorliegenden Studien zeigten insgesamt keine konsistente Konkordanz der immunhistochemischen Bestimmungsergebnisse für Her-2/neu aus Biopsiematerial im Vergleich zum definitiven Operationspräparat. Die beste Übereinstimmung wurde für die Kategorie IHC3+ nachgewiesen.

Für die Methode FISH liegt nur aus einer Studie das Ergebnis einer vollständigen Übereinstimmung vor.

Bei der Wahl der Bestimmungsmethode am Biopsiematerial ist zusätzlich zu der inkonsistenten Konkordanz die grundsätzlich schlechtere Reliabilität der IHC zu berücksichtigen. Derzeit kann die immunhistochemische Bestimmung aus Biopsiematerial aufgrund der nicht sicheren Übereinstimmung mit dem definitiven Operationspräparat nicht empfohlen werden.

Weitere prospektive Studien mit verblindeter Beurteilung der Gewebeproben sind für eine valide Beurteilung der Übereinstimmung von Biopsieergebnis und Ergebnis am definitiven Operationspräparat für beide der Bestimmungsmethoden erforderlich.

5. Evidenztabellen

5.1 Ergebnisse und Evidenzgrad von IHC und FISH in Bezug auf das Ansprechen auf Trastuzumab

5.1.1 Eingeschlossene Studien

Tabelle 1

Ergebnisse und Evidenzgrad von IHC und FISH zum Therapieansprechen auf Trastuzumab

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Hofmann M. et al., 2007, BRD	Retrospektiv IHC und FISH aus Archivmaterial von Phase II Studie WO16229 (first line Herceptin Mono Therapie)	Sensitivität FISH+: 82,4% Sensitivität IHC3+: 100% (CR = 2, PR= 17, : alle IHC3+ , 3 FISH- 15,7%).	Bei Polysomie wird IHC3+ als für Therapieansprechen entscheidend gesehen	3b
	Studiengruppe: n= 95 von 103/105 mit Response Daten von insgesamt 289 Pat.	27% Polysomie, 6 Response auf Trastuzumab, alle IHC3+, 2x FISH - wegen der Polysomie.		
	Zeitraum: -			
	Outcome Parameter: 1. Overall Response 2. Time to progression			
	Gewebeproben: Paraffinschnitte			
	Fixierung: zu > 90% Formalin, n.n.def.			
	Test: Hercep Test, Pathvysion			
	Scoring-System: n. DAKO, FISH+ >/=2.0, Polysomie >/= 3 CEP 17 Kopien pro Zelle			
	Verblindeter Vergleich: überwiegend			
Mitgelaufene Kontrollen: ja				
Mass RD et al., 2005, USA	Retrospektiv IHC und FISH aus Archivmaterial	FISH +: 78% FISH -: 22% (72% IHC2+)	Keine Zahlenangaben über Response IHC 3+ allein	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	<p>Studiengruppe: n= 784 von 799 Pat. in 3 klinischen Studien zur Behandlung von met. Brustkrebs mit Trastuzumab mit IHC 2+ und 3+, davon n= 765 mit FISH (95,7%)</p> <p>Zeitraum: -</p> <p>Outcome Parameter: 1. Overall Response 2. Time to progression 3. Overall Survival</p> <p>Gewebeproben: Paraffinschnitte</p> <p>Fixierung: nicht angegeben</p> <p>Test: IHC: CTA = monoklonaler AK 4D5 oder monoklonaler AK CB11 FISH: dc PathVysion</p> <p>Scoring-System: IHC: 0-3, hier nur 2+ und 3+, FISH: >/=2 Ratio ist positiv, <2 negativ</p> <p>Verblindeter Vergleich: nicht angegeben</p> <p>Mitgelaufene Kontrollen: nicht angegeben</p>	<p>28% IHC 3+)</p> <p>Studie 1 (H06489g – a)Trastuzumab +Chemotherapie versus b = Chemo alleine):</p> <p>1. Overall Response: FISH+: a)54% b) 30% p=0.001 FISH-: a) 38% b) 39%</p> <p>2. Time to progression: FISH+: RR 0.45 p<0,0001 FISH-: RR 0.61 p=0.03</p> <p>3. Overall Survival: FISH+: RR 0.71 p=0.009 FISH-: RR1.10</p> <p>Studie 2 (H0649g – single Arm-Behandlung mit Trastuzumab nach 1-2 Chemotherapien:</p> <p>1. Overall Response: FISH+: 19% FISH-: 0%</p> <p>2. Time to progression : überlappend</p> <p>3. Overall Survival: überlappend</p> <p>Studie 3 (H0650g – randomisiert Trastuzumabschema 1 oder 2):</p> <p>1. Overall Response: FISH+: 34% FISH-: 7%</p> <p>2. Time to progression: FISH+: 4,9 Monate FISH-: 1,7</p> <p>3. Overall Survival: gleich</p>		
Seidman A.D. et al., 2001, USA	<p>Prospektiv, Phase II Studie</p> <p>Studiengruppe: alle n= 95 Pat. mit metast. Mammakarzinom und bis zu 3 vorherigen Chemotherapien, Behandlung mit Trastuzumab und Paclitaxel mit und ohne Her2-Überexpression</p>	<p>Failure FISH: n= 16</p> <p>Korrelation Overall Response mit Her2-Ergebnissen:</p> <p>- + FISH: 44% 75% p=0.004 HercepTest: 46% 69% p=0.03</p>	Keine absolute Zahl zu 3+ alleine	2b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Zeitraum: - Outcome Parameter: Overall Response Gewebeprobe: Paraffinschnitte Fixierung: Formalin, nicht näher angegeben Test: IHC: DAKO HercepTest Pab 1 TAB250 Cb11 FISH: dc PathVysion Scoring-System: nach DAKO FISH: >/= 2 Ratio ist positiv Verblindeter Vergleich: ja und Beurteilung durch 2 Pathologen Mitgelaufene Kontrollen: -	Pab1: 41% 67% p=0.02 Cb11: 45% 76% p=0.004 TAB250 43% 81% p=0.001 Alle statistisch signifikant Herceptest zeigt im Vergleich zu CB11 9 mehr 2+/3+: Konkordanz zu FISH+ CB11: 2+/3+: 34/36 94,4% HercepT.: 2+/3+: 37/45 82,2%		
Vogel C.L. et al., 2002, USA	Prospektive Kohortenstudie, FISH aus Archivmaterial Studiengruppe: n= 111 Pat. mit metast. Brustkrebs und IHC 2+ oder 3+ in Phase 3 Studie randomisiert, First line Trastuzumab in 2 Dosierungen und FISH Zeitraum: - Outcome Parameter: Primär: -Overall Response Sekundär: 1. Time to progression 2. Survival 3. Quality of life (Ergebnisse nicht dargestellt) Gewebeprobe: FFPS Fixierung: nicht angegeben Test: IHC: CTA (4D5 und CB1)1 FISH: Pathvysion Scoring-System: IHC 2+ moderate Anfärbung >10% der Zellen, IHC 3+ starke Anfärbung >10% der Zellen, FISH: nicht angegeben Verblindeter Vergleich: - Mitgelaufene Kontrollen: -	Response IHC 3+: (n=84): 35% Response IHC 2+: (n=27): 0% Response FISH +: (n= 79): 34% Response FISH - : (n= 29): 7%		3b

5.1.2 Ausgeschlossene Studien

Tabelle 2

Ausgeschlossene Studien zum Therapieansprechen auf Trastuzumab

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Burststein et al., 2003, USA	First line Trastuzumab und Vinorelbine n= 54, n= 44 mit IHC 3+ und n= 10 mit FISH+. Response-Raten für beide Methoden gleich	Keine vergleichende Aussagen möglich, da jeweils nur eine Bestimmungsmethode
Diaz et al., 2001, USA	Review zum Ansprechen von Trastuzumab unter Berücksichtigung der Mass und Vogel-Studie	Keine Originaldaten
Fornier et al., 2005, USA	Retrospektive Analyse für Serum Her2 , Vergleich IHC und FISH	Originaldaten aus Seidman et al
Hudelist et al., 2004, USA	Review zur Trastuzumab-Therapie mit Diskussion der Bestimmungsmethoden	Keine Originaldaten
Jahanzeb et al., 2003, USA	Review zur Trastuzumab-Therapie mit Diskussion der Bestimmungsmethoden	Keine Originaldaten
Leyland-Jones , 2001, USA	Review	Keine Originaldaten
Pegram et al., 2004, USA	Retrospektive Fish-Analyse an Pat. mit met. Karzinom und IHC2+, IHC3+ oder FISH + lokal a) n= 56 v 62 Pat. aus BCIRG 101-Studie (Trastuzumab, Docetaxel und Cisplatin) zeigt keinen Unterschied in Response bei FISH+ und FISH- b) n= 57 v. 59 Pat. aus UCLA-ORN-Studie zeigt höhere Response-Rate bei FISH+ als bei FISH-, stat. n. sign.	FISH-Daten können den IHC'/FISH lokal-Daten nicht zugeordnet werden
Tedesco et al., 2004, USA	Prospektiv, Phase II Studie Pat. mit met. Mammakarzinom und first line Docetaxel + Trastuzumab mit IHC und FISH, Response 1x bei FISH positiv und IHC2+ und 1x bei FISH negativ (IHC-Status unklar) alle IHC3+ FISH+, Ansprechrate IHC3+ 63%, FISH+ 64%, IHC2+ 14%, FISH- 25% (? - unklar)	Ergebnisse im Hinblick auf Vergleich IHC-FISH nicht eindeutig dargestellt.
Vogel et al., 2001, USA	Review	Keine Originaldaten

5.2. Ergebnisse und Evidenzgrad zur Reliabilität von IHC und FISH

Tabelle 3

Ergebnisse und Evidenzgrad zur Reliabilität von IHC und FISH

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Bartlett et al., 2001, UK	Retrospektive Kohortenstudie	Reliabilität (Interobserverbeurteilung)	Radioimmunohistochemie als Goldstandard nicht akzeptiert, Ergebnisse nicht dargestellt	3b
	Studiengruppe: n= 216	Herceptest: 93%		
	Zeitraum: 1984-1998	CB 11: 62%		
	Gewebeproben: ‚frozen Sections‘	FISH: 95%		
	Fixierung: Formalin n.n. definiert			
	Test: -HercepTest DAKO -CB 11 mit Antigen Retrieval, FISH und Radioimmunhistochemie (Goldstandard)			
	Scoring-System: n. DAKO 2+ und 3+ als positiv gewertet und > 2.0 Ratio Fish +			
	Verblindeter Vergleich: ja, 2 unabhängige Befunder			
Mitgelaufene Kontrollen: -				
Bilous et al., 2002, Australien	Kohortenstudie aus 13 Laboren, zum Teil prospektiv (als Teil der Her2000 Studie)	IHC-Interobserver-Reliabilität (14 Pathologen aus 13 Laboren beurteilen 5 IHC-slides): K= 0.44		3b
	Studiengruppe: n= 1.536 mit Herceptest davon n= 109 mit FISH bei IHC 2+			
	Zeitraum: -			
	Gewebeproben: Paraffinschnitte			
	Fixierung: gepuffertes Formalin n.n.def.			
	Test: -HercepTest (manuell oder automatisch) -PathVysion			
	Scoring-System: IHC n.DAKO, 2+ und 3+ als positiv, FISH+ >2.0 Ratio			
	Verblindeter Vergleich: -			
Mitgelaufene Kontrollen: ja				

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Dowsett et al., 2003, UK	Retrospektive Kohortenstudie Studiengruppe: n= 60 von n=426 von 486 mit met. Mammakarzinom im EAP Programm für kontrollierten Einsatz von Trastuzumab aus 37 Laboren Bestimmungen in 3 Referenzlaboren FISH und IHC Zeitraum: 1-6/2001 Gewebeprobe: Paraffinschnitte Fixierung: Formalin, n.n. def. Test: HercepTest und PathVysion Scoring-System: n. DAKO und >/= 2.0 Ratio+ Verblindeter Vergleich: - Mitgelaufene Kontrollen: -	IHC-Reliabilität Interlabor (3 Referenzzentren): 66% n= 60 10+ diskordant 2+/3+: alle FISH+ 10+ diskordant 1+/2+: 9 FISH- 1x FISH + (3.15)		3b
Dowsett et al., 2007, UK	Prospektive Vergleichsstudie Studiengruppe: n=20 IHC und 20 FISH in 5 internationalen Referenzlaboren (London, Toronto, Antwerpen, Kassel, Amsterdam) Zeitraum: 5 Testrunden im Abstand von etwa 2 Monaten Gewebeprobe: Paraffinschnitte Fixierung: Formalin n.n. definiert (12-48h) Test: -HercepTest DAKO - Pathvysion, Vysis Scoring-System: n. DAKO und >/= 2.0 Ratio Fish + Verblindeter Vergleich: ja Mitgelaufene Kontrollen: ja	Interlabor(5)- Reliabilität IHC: 45% (9/20) 8x mind. 1 v. 5 Zentren 0/1+ und die anderen 2+ 3x mind. 1 v.5 Zentren 3+ und die anderen 2+ FISH (2/11 technical failure) 5/9 negativ in allen Zentren, 1/9 positiv in allen Zentren 3 diskordant Interlabor(5)-Reliabilität FISH: 80% (16/20) 4 diskordante Fälle alle mit mind. 1 Befund von 1.7-2.3	Betonen die Notwendigkeit rigoroser Qualitätskontrollen	3b
Dybdal et al., 2005, USA	Bestimmung von FISH aus Archivmaterial mit bereits bestimmtem IHC durch CTA Studiengruppe: n= 623 in 2 Laboren Zeitraum: - Gewebeprobe: Paraffinschnitte. Fixierung: n. n. def.	Reliabilität FISH (Interlabor 2) n=488 92% K=0.79 37x FISH- in Labor 1 und FISH+ in Labor2: davon 35 IHC3+!		3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Test: - CTA <i>FISH: PAthvysion (mit Protokollmodifikationen)</i> Scoring-System: 0-3 (10%, 1 = inkomplett, 2= komplett schwach, 3= komplett stark) <i>FISH+Ratio >/= 2.0</i> Verblindeter Vergleich: ja Mitgelaufene Kontrollen: +/- Zelllinien	LA (1.8-2.2) <2%		
Gancberg et al., 2002, Belgien	prospektive Kohortenstudie Studiengruppe: n= 394 in 2 Laboratorien (in Brüssel und in Tampere, Finnland) beide mit monoklonalem AK CB11 aber unterschiedlichen Protokoll und unterschiedlichem Scoring-System für IHC getestet, n= 254 auch mit FISH (in Brüssel durchgeführt) Zeitraum: Gewebeprobe: Paraffinschnitte	Interlabor-Reliabilität (2) gesamt: 84,5% Reliabilität +: 100% Reliabilität -: 82,4% Diskordanz höher in Formalin-Fixierter Gruppe: <u>Her-2 positiv:</u> Brüssel: 12,2% Tampere: 27,7% p<0,001	FISH nicht möglich bei Bouin-Fixierung Trotz konservativem Scoring in Tampere signifikant höhere + Rate: Grund vermutet in Antigen Retrieval: erhöht Sensitivität und erniedrigt Spezifität	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	<p>Scoring-System: IHC Brüssel: 0=negativ: keine angefärbten Zellen 1+= >1% und <25% 2+= >25% und <50% 3+= >50% und < 75% 4+= >75%</p> <p>Cut-off positiv: >1%!</p> <p>Tampere: a.negativ: Membranfärbung in < 50% und nicht intensiv b. positiv: > 50%, intensive Färbung</p> <p>FISH: Nicht amplifiziert: Ratio </=2 LA: >2, </=5 HA: >5</p> <p>Verblindeter Vergleich: ja</p> <p>Mitgelaufene Kontrollen: in Tampere Tumorgewebe und Zelllinien mit bekannter positiver oder negativer Überexpression <i>Goldstandard</i> FISH</p>	<p>Gen-Amplifikation FISH: 22,2% (FISH n=248 auswertbar) 20% versus 7% Bouin</p> <p>Brüssel Tampere</p> <p>NPV IHC: 0.85 0.98 PPV IHC: 0.88 0.68</p> <p>Rate falscher Ergebnisse: 0.15 0.11</p>		
Hsu et al., 2002, Taiwan	<p>Prospektive Kohortenstudie Vergleich Beurteilung durch 5 Pathologen</p> <p>Studiengruppe: n= 46 IHC DAKO beurteilt von 5 Pathologen</p> <p>Zeitraum: 1988-1989</p> <p>Gewebeproben: nicht angegeben</p> <p>Fixierung: nicht angegeben</p> <p>Test: Herceptest, DAKO, stand. Durchführung</p> <p>Scoring-System: <i>lt. Herceptest</i></p> <p>Verblindeter Vergleich: ja</p>	<p>Interobserver Reliabilität aller 5 Pathologen: 48% K=0.8</p> <p>Relative Reliabilität (+/-) : 95,6% K=0.96</p> <p>Reliabilität Beurteilung 2+/3+ n=22: 59% K=0.38</p> <p>Reliabilität : Membranfärbung vorhanden: K=0.82 Starke Anfärbung: K=0.8</p>	Autoren fordern Einhaltung des HercepTest und mitlaufende Kontrollen, positive Fälle sollen mit FISH getestet werden	

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Mitgelaufene Kontrollen: normales Gewebe	,Completeness' K=0.96 Anteil angefärbter Zellen K=0.7 Hoher NPV – guter Ausschluss negativer Fälle		
Lacroix-Triki, M et al. (GEFPICS), 2006, Frankreich	Prospektive Kohortenstudie Studiengruppe: n= 74 slides, beurteilt von 16 Pathologen und von einem zentralen Komitee Zeitraum: 2003 Gewebeproben: Paraffinschnitte Fixierung: neutral gepuffertes Formalin n= 43 Alkohol-Formalin-Essig-Säure n= 19 Hollands Bouin n= 9 Bouin Fixierung n= 4 Test: CB 11, Novocastra Newcastle n= 39, A0485, DAKO, Glostrup n= 36 <i>Besonderheiten der Durchführung: 11 von 12 Laboren führten ein Antigen-Retrieval mit Citratpuffer mit verschiedenen Methoden durch Detektion biotinlabelled außer in einem Labor: Dextran-Polymer (Envision, DAKO) Unterschiedliche Antikörperverdünnung und Inkubationszeit</i> Scoring-System: - Verblindeter Vergleich: ? Mitgelaufene Kontrollen: nein	Interlabor-Reliabilität IHC 0/1: K= 0,85 IHC 2: K= 0,38 IHC 3: K= 0,82 Overall: K=0.75 Overall bei +/-: K=0.85 Untergruppe IHC 1 (n=10 durch Experten): In fast 50% der Befundungen durch Einzelbefunder als 2+ gescort. K=0.44	Kein externer Referenzstandard, sondern Expertenkonsensus als Standard genutzt Große Differenz bei 1+/2+ und 2+ sehr heterogene Gruppe: 4 Fixierungen, 2 Tests etc	3b
Paik et al., 2002, USA	retrospektive Kohortenstudie Studiengruppe: Vergleich n= 104 IHC 3+ aus 67 lokalen Laboren (konsekutiv seit Beginn der Studie) mit zentralem Labor im Rahmen der NSABP B-31-Studie Zeitraum: 2001	Interlabor-Reliabilität: gesamt 79% (n= 22 different) FISH positiv: 79% Zentral werden 19 Fälle als negativ bewertet (nicht stark positiv für IHC oder amplifiziert mit FISH), davon	Da aus vielen kleinen Laboren nur 1 Slide kam, kann Differenz überschätzt werden	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Gewebeproben: <i>Paraffinschnitte</i> Fixierung: <i>nicht angegeben</i> Test: <i>1. lokal HercepTest (DAKO) n= 80 = 77% Andere n= 24 2. zentral Herceptest (DAKO) PathVysion (FISH)</i> Scoring-System: <i>Vom Hersteller</i> Verblindeter Vergleich: - Mitgelaufene Kontrollen: -	n= 10 Score 0 oder 1+ n= 9 Score 2+ Reliabilität bezogen auf Laborgröße (cut-off 100 Fälle pro Jahr). 76% Reliabilität für kleine Labore n= 58 (differente Fälle häufiger bei Nicht-Anwendung von Hercep-Test) 97% Reliabilität für große Labore n= 9 (diese benutzten nur HercepTest) Konkordanz Zentrallabor FISH-IHC: 98/104 = 94,3% Interlabor-Reliabilität FISH (validiert in einem zweiten Labor n= 81) 95%		
Perez et al., 2006, USA	retrospektive Kohortenstudie Studiengruppe: n= 2535, seit 3/2002 rekrutiert für den N9831 Intergroup Trail mit positivem Her-2/neu durch IHC (3+) oder FISH (>/= 2.0) Zeitraum: 3/2002-4/2005 Gewebeproben: FFPS Fixierung: nicht angegeben Test: IHC lokales Labor: 42,1% HercepTest 25,2% andere FISH lokales Labor: 32,7% PathVysion IHC zentrales oder Referenzlabor: HercepTest FISH: PathVision Scoring-System: Vom Hersteller HercepTest FISH cut-off 2 Verblindeter Vergleich: -	Technisches Versagen FISH: 3,1% Technisches Versagen IHC: 0,4% 1. Interlabor-Reliabilität lokal-zentral: Reliabilität Overall: 85.5% Reliabilität bei gleicher Methode: a. IHC Herceptest: 81,6% b. IHC andere AK: 75% c. FISH: 88,1% p<0,001 signifikant besser für FISH 2. Interlabor-Reliabilität zentral-Referenz Reliabilität overall: 90,7% Reliabilität bei gleicher Methode a. IHC (n= 300): 94,3% 17x wurde das zentral 0,1+ oder 2+		3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Mitgelaufene Kontrollen: ja, im zentralen und Referenzlabor positive und negative Kontrollen aus HercepTest	vom Referenzlabor in 3+ upgegraded. b.FISH (271 von n=303 auswertbar):95,2% 13x zeigte sich FISH-Amplifikation im Referenzlabor		
Press et al., 2005, USA	<p>prospektive Kohortenstudie</p> <p>Studiengruppe: <i>n= 2.502 mit FISH in 2 zentralen Laboren (Basel und Los Angeles) im Rahmen von 3 BCIRG Studien (005,006,007) von n= 2.600 konsekutiven, Vergleich mit n= 1.617 mit IHC oder FISH lokal</i></p> <p>Zeitraum: -</p> <p>Gewebeproben: <i>Paraffinschnitte</i></p> <p>Fixierung: <i>nicht angegeben</i></p> <p>Test: <i>zentral: FISH: dc PathVision, Abbott-Vysis IHC: 10H8-monoklonaler AK Lokal: DAKO HercepTest PathWay, Ventana andere</i></p> <p>Scoring-System: <i>FISH: <2 negativ />= 2 positiv</i></p> <p>Verblindeter Vergleich: ja</p> <p>Mitgelaufene Kontrollen: -</p>	<p>Interlabor-Reliabilität 2 zentrale Labore: FISH (n=50, nicht in Studie): 98,7% K=0.97</p> <p>FISH Interobserver-Reliabilität (n= 59, nicht in Studie): 96,6% K=0.93</p> <p>FISH+: 26,3%, technical failure ca.1,6%</p> <p><u>Konkordanz IHC lokal-FISH zentral: 79%</u> K= 0.56</p> <p>Reliabilität IHC lokal-zentral n=1.291 77,5% K=0,51</p> <p>Reliabilität FISH lokal-zentral n=1.311 : 92% K= 0.83</p>	FISH weitaus objektiver	3b
Reddy et al., 2006, USA	<p>Prospektive Kohortenstudie</p> <p>Studiengruppe: <i>n= 973 aus n=1503 aus HER-First-Studie, Proben aus lokalen Laboren, die zentral nachgetestet wurden, Einschluss in Studie: 3+ oder amplifiziert</i></p> <p>Zeitraum: <i>1/2001-2/2004</i></p> <p>Gewebeproben: <i>Paraffinschnitte</i></p> <p>Fixierung: <i>nicht angegeben</i></p> <p>Test: <i>lokal: nicht angegeben zentral: HercepTest</i></p> <p>Scoring-System: <i>0-3 lt. Hersteller, 3+ wird als positiv gewertet</i></p>	<p>Interlabor-Reliabilität IHC lokal-zentral (n= 973): Score 0: 34% Score 1: 54% Score 2: 26% Score 3: 77%</p>		3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Verblindeter Vergleich: - Mitgelaufene - Kontrollen:			
Rhodes et al., 2002, UK	Prospektive Kohortenstudie Studiengruppe: n= 4 Zelllinien werden 2x von 78 (von insgesamt 94) Laboren beurteilt, vergleichende Auswertung durch 5 externe Pathologen und n=105 Mammakarzinome Zeitraum: - Gewebeprobe: FFPS Fixierung: neutral gepuffertes Formalin Test: - Scoring-System: das für CTA benutzte Verblindeter Vergleich: ? Mitgelaufene - Kontrollen:	IHC Score Zelllinien wurden von den 5 Bewertern festgelegt: a) 3+, b)2+, c)1+, d)0. IHC-Interlabor-Reliabilität gesamt (2.Runde): HercepTest 76% A0485 55% CB11 40% Andere (n= 7) 71% Objektivität gesamt war in der 2 Runde signifikant besser für A0485 und CB11 im Vergleich zur 1 Runde Reliabilität Zelllinien gesamt: 74% Reliabilität Zelllinien HercepTest (auch Scoring-system) (26 Labore): 86%. K= 0.81	HercepTest wird empfohlen. Auch für andere AK kann die vergleichende Nutzung von standardisierten Zelllinien Qualität verbessern	3b
Roche et al., 2002, USA	retrospektive Kohortenstudie Studiengruppe: n= 119 (erste konsekutive Fälle aus dem Intergroup Trial N9831), Rekrutierung von Pat. mit Her2-neu Überexpression (3+ oder amplifiziert) Zeitraum: Gewebeprobe: Paraffinschnitte Fixierung: nicht angegeben Test: lokal IHC: HercepTest 50%, andere 42% FISH: PathVysion o. Ventana (ins.7%) zentral IHC: HercepTest mit DAKO Autostainer/ Fish: PathVysion	IHC Interlabor-Reliabilität lokal und zentral: 74% Reliabilität IHC 3+ n= 75% Reliabilität FISH lokal zentral = 67% (6/9) Konkordanz IHC-Fish zentral: 92% (2+ wird als negativ bewertet)	Aufgrund der Ergebnisse wurde die Aufnahme in die Studie von Bestätigung des Befunds durch ein zentrales Labor abhängig gemacht	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Scoring-System: <i>IHC: >33% der Zellen stark positiv FISH: > 2.0 oder > 5 Gene Kopies</i> Verblindeter Vergleich: ? Mitgelaufene Kontrollen: -			
Tsuda et al., 2001, Japan	Prospektive Kohortenstudie, daraus Ergebnisse für Reliabilität FISH dargestellt Studiengruppe: n= 216 , konsekutiv beurteilt für FISH von 2 Befundern Zeitraum: <i>1999-2000</i> Gewebeproben: <i>FFPS</i> Fixierung: <i>10% und 20% Formalin, nicht näher definiert</i> Test: <i>dc Pathvysion</i> Scoring-System: <i>>/= 2 Signale positiv</i> Verblindeter Vergleich: ja Mitgelaufene Kontrollen: <i>ja, positive und negative</i>	Interobserver-Reliabilität FISH: Reliabilität Her2/CEP 17-Ratio: 99% Stratifizierung in LA und HA: 98% K=0.94 Anzahl Her2+ = 19%	Ergebnisse Konkordanz IHC-Fish in dieser Tabelle nicht dargestellt.	2b
Tsuda, 2002, Japan	Studiengruppe: n= 106 beurteilt von 6 bzw 5 Pathologen Zeitraum: <i>1990-1995</i> Gewebeproben: <i>FFPS</i> Fixierung: <i>Formalin, nicht näher definiert</i> Test: <i>Herceptest und jap. polyklonaler AK (n=50) sowie CB11 und TAB250 (jew. n= 56)</i> Scoring-System: <i>0-3 nach DAKO und jap. System (von diesem sind Ergebnisse nicht dargestellt)</i> Verblindeter Vergleich: ja	Interobserver-Reliabilität IHC: 1. bei 2+und3+ als+ gewertet: HercepTest: 84% Jap. AK: - CB11: 93% TAB250: 95% 2. bei 3+ als positiv gewertet HercepTest: 94% (sign. besser p=0,011) Jap. AK: - CB11: 89% TAB250: 96% Vergleich Her2 3+		3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Mitgelaufene Kontrollen: -	HercepTest: 12% Jap. AK: - CB11: 5,3% TAB250: 5,3%		
Yaziji et al., 2004, USA	retrospektive Kohortenstudie Studiengruppe: n= 117 mit IHC zentral von n=2.963 ges. von 135 Laboren Zeitraum: 1/1999-5/2003 Gewebeprobe: Paraffinschnitte Fixierung: verschiedene Fixierungen, n.angegeb (wide range). Test: - A0485 DAKO mit AG Retrieval (Citratpuffer) - FISH von Vysis Scoring-System: - IHC: ‚subtraction score‘: tumorzellscore-Hintergrundscore 0-3 - FISH+ LA Ratio >/= 2.0-<5; HA >/=5.0 Verblindeter Vergleich: - Mitgelaufene Kontrollen: ja	Reliabilität IHC 5 Interobserver: (absolut) 50,4% Reliabilität IHC 4 Observer: (absolut) 75,1% keine ‚two-step‘-Abweichung (nur 0-1+ und 2+-3+)		3b
Thomson et al., 2001, Kanada	prospektive Kohortenstudie Vergleich Studiengruppe: <i>n= 95 aus 24 Laboren daraus, IHC (4 AK) Beurteilung durch 3 unabhängige Pathologen und FISH</i> Zeitraum: 1/-12/1999 Gewebeprobe: Paraffinschnitte Fixierung: 10% gepuffertes Formalin Test: IHC: DAKO A0485, Novocastra CB11, ZymedTAB250, DAKO HercepTest, FISH: dc Vysis Besonderheiten der Durchführung: Antigen.Retrieval nur fürTAB250 und Herceptest Ventana ES Immunostainer	Interlabor-Reliabilität: Overall +/- (2+/3+==+) A0485 : K=64,2 79,5 % CB11: K=74,4 96,8% TAB250: K=79,2 86,0 % HercepTest: K= 69,5 81,6 % Reliabilität für 0 und 3+ : K= 0,77-0,96 Reliabilität für 1+ und 2+: K= 0,33-0,59	Fixierungsdauer für die 4 Ak variierte beträchtlich, Schnitte waren bis zu 13 Jahre alt.	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Scoring-System: <i>IHC: 0-3 nach DAKO, 2+ und 3+ als positiv gewertet, FISH cut-off >/=2</i>			
	Verblindeter Vergleich: ja			
	Mitgelaufene Kontrollen: ja, positive und negative			

5.2.1 Ausgeschlossene Studien zur Reliabilität von IHC und FISH

Tabelle 4

Ausgeschlossene Studien zur Reliabilität von IHC und FISH

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Günhan, 2004, Türkei	Interlabor-Vergleich von 2 IHC-AK	Einer der AK ist hier nicht gebräuchlich (Clone e2-4001+3B5)
INQAT Group, 2004, Italien	Qualitätsprogramm IHC Her-2neu-Bestimmung – Intra- und Interobserverbeurteilung	Verwendete AK nicht angegeben

5.3. Ergebnisse und Evidenzgrad Konkordanz IHC und FISH

Anmerkung:

In Tabelle 5 sind Studien gelistet, in denen eine Stratifizierung von 0-3+ vorgenommen wurde, dabei sind zunächst alle Studien mit Verwendung des Herceptests gelistet.

Tabelle 6 enthält Studien, die nur nach +/- diskriminieren oder nur eine IHC-Gruppe analysiert haben (2+ oder 3+).

Tabelle 5

Ergebnisse und Evidenzgrad Konkordanz IHC und FISH -I- stratifiziert nach 0/1+, 2+, 3+

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Ainsworth et al., 2005, Irland	Retrospektive Kohortenstudie, Bestimmung aus Archivmaterial	Konkordanz IHC-FISH:	IHC 2+ und 3+ sehr kleine Gruppengrößen, Ergebnisse statistisch nicht für eine signifikante Differenz gepowert Gesamtüberleben: in allen Gruppen vergleichbare Ergebnisse bei cut-off IHC zwischen 1+ und 2+ ! jeweils statistisch signifikant	3b
	Studiengruppe: n= 165 mit invas. Karzinom, med. Alter 70 J.	HercepTest: 2+ (12/13): 92% HercepTest 3+ (<u>11/12</u>): 92%		
	Zeitraum: 1984-1998, med. Follow-up 5 J.	In Auswertung nicht eingerechnet, da nicht gerbäuchlich (CBE356 2+ (6/10) 60% CBE356 3+ (<u>26/28</u>) 93%		
	Gewebeprobe: Paraffinschnitte			
	Fixierung: Formalin, n.n. def.			
	Test: - HercepTest mit automat.Staining n. DAKO			
	- monoklonaler AK CBE356, Novocastra (ext. Domäne) mit AG Retrieval (EDTA)			
	- dc PathVysion (Goldstandard)			
	Scoring-System: IHC n. DAKO, FISH + > Ratio 2.0			
Verblindeter Vergleich: -				
Mitgelaufene Kontrollen: ja (von HercepTest)				
Outcome-analyse: Gesamtüberleben				
Bouchalova et al., 2006,	Prospektiv Kohortenstudie Bestimmung Her2 und TOP2a	Technical Failure FISH: 0,4% (n= 3)	Ergebnisse für TOP2a nicht dargestellt	3b
	Studiengruppe: n=72 randomisiert ausgewählt	Konkordanz IHC-FISH:		
	Zeitraum: 4/1998-4/2000			

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Tschechien	Gewebeproben: Paraffinschnitte Fixierung: 10% gepuffertes Formalin Test: - Herceptest DAKO - three colour, Vysis (mit TOP2a) Scoring-System: IHC n. DAKO FISH+ >/= 1.5 Ratio Verblindeter Vergleich: - Mitgelaufene Kontrollen: -	0/1+ (n= 47) 100% (alle FISH-) 2+ (n= 12) 33% (8 FISH-) 3+ (n= 10) 100% (alle FISH+)		
Ciampa et al., 2006, USA	Retrospektive Kohortenstudie Studiengruppe: n= 108 Zeitraum: - Gewebeproben: <i>Paraffinschnitte</i> Fixierung: Formalin 10% gepuffert Test: - HercepTest nach DAKO Mit ACIS automated cellular imaging system, Chroma Vision) - dc PathVysion Scoring-System: n. DAKO und ACIS (+=ab2.6) Verblindeter Vergleich: - Mitgelaufene Kontrollen: -	Konkordanz IHC-FISH: Herceptest n.DAKO: 0/1+ n= 21/23: 91% 2+/3+ n= 25/85: 30,5% if 2+ als negativ bewertet:- 2+ n= 40/43: 93% 3+ n= 22/42: 52% Herceptest n. ACIS </= 1.5 n= 22/23: 96% 1.6-1.9 n= 15/17: 88% 2.0-2.5 n= 41/46: 89% >/= 2.6 n= 19/22: 86%	Empfehlen ACIS (+Score) und FISH für 1.6-2.5	3b
Dolan, Snover 2005, USA *	Retrospektive Kohortenstudie Studiengruppe: n= 129 von 275 im Zeitraum mit FISH im Referenzlabor und IHC aus mehreren Zusenderlaboren davon 42 core biopsy Zeitraum: 10/01-3/04 Gewebeproben: Paraffinschnitte Fixierung: Formalin, n.n.def. Test: HercepTest DAKOCytomation Dc PathVysion Vysis Scoring-System: n. DAKO FISH Ratio >/=2.0 + Verblindeter Vergleich: -	Konkordanz IHC-FISH: 0/1+ = 17,8% n= 23: 100% (alle FISH-) 2+ = 72,1% n= 93: 8% (7x FISH+) 3+ = 10,1% n= 13: 38% (5x FISH+) Konkordanz FISH-IHC: FISH+: n= 12: 100% (alle IHC 2+ oder 3+) FISH-: n= 117: 19,7% (20,3% IHC 2+, 3+) Konkordanz IHC-FISH bei CORE Biopsy war		3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Mitgelaufene - Kontrollen:	schlechter, hier konnten nicht durchgängig 60 Zellen(Zellkerne) beurteilt werden		
Dowsett et al., 2003, UK *	Retrospektive Kohortenstudie Studiengruppe: n= 426 von 486 mit met. Mammakarzinom im EAP Programm für kontrollierten Einsatz von Trastuzumab aus 37 Laboren Bestimmungen in 3 Referenzlaboren FISH und IHC Zeitraum: 1-6/2001 Gewebeproben: Paraffinschnitte Fixierung: Formalin, n.n.def. Test: HercepTest und PathVysion Scoring-System: <i>n. DAKO und >/= 2.0 Ratio+</i> Verblindeter Vergleich: - Mitgelaufene - Kontrollen:	Konkordanz IHC-FISH: 0/1+ n= 268/270: 99,3% 2+ n= 26/54: 48,1% 3+ n= 96/102: 94,1% Reliabilität Interlabor (3 Zentren): 66% n= 60 10+ diskordant 2+/3+: alle FISH+ 10+ diskordant 1+/2+: 9 FISH- 1x FISH + (3.15)		3b
Ellis et al., 2005, UK	Retrospektive Studie Studiengruppe: n=114 aus Archivmaterial Zeitraum: - Gewebeproben: <i>Paraffinschnitte</i> Fixierung: Formalin n.n.def. Test: HercepTest, DakoCytomation und mit PathVysion (manuell und mit MDS) Scoring-System: <i>It Hersteller, FISH+ > 2.2 Ratio</i> Verblindeter Vergleich: - Mitgelaufene - Kontrollen:	Konkordanz IHC-FISH manuell: 0/1+ = n= 66/72: 92% 2+ = n= 8/22: 36% 3+ = n= 18/20: 90% 2x Borderline amplifikation (1.8-2.2) FISH (IHC1+ und IHC2+) Konkordanz IHC-FISH automatisch: Unterschied in 21 Fällen (18%). 14x 2+ manuell werden 1+ oder 3+ gescort und 13 davon sind dann konkordant mit FISH: d.h. höhere Konkordanz	Empfehlen automatisches System	3b
Hanna et al., 2006, USA	Prospektive Bestimmung IHC, FISH/CISH Studiengruppe: n= 254 , unklar, ob konsekutiv Zeitraum: - Gewebeproben: FFPS Fixierung: -	Failure: 0% Konkordanz IHC und FISHdc: 0/1+: (n= 69): 100% 2+: (n= 37/135): 27,4%		3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Test: IHC = Herceptest, DAKO , FISH = PathVysion dual colour nach Angaben der Hersteller IHC n. DAKO Scoring-System: FISH Ratio >/02.0 + Verblindeter Vergleich: ja Mitgelaufene Ja Kontrollen:	3+: (n= 46/50) 92%		
Hoang et al., 2000, USA *	Retrospektive Kohortenstudie Studiengruppe: n= 100 konsekutiv Zeitraum: - Gewebeprobe: Paraffinschnitte Fixierung: Formalin n.n.def. Test: -Herceptest, - e2-4001, Neomarker (Ergebnisse nicht dargestellt) - PAtHvysion dc Scoring-System: lt Hersteller und FISH+ Ratio >/= 2.0 Verblindeter Vergleich: - Mitgelaufene nach Kit Kontrollen:	Konkordanz IHC und FISH: 0/1+ n= 74/74 100% 2+ n= 1/8 12,5% 3+ n= 16/18 89%	Ergebniss für e3-4001 nicht dargestellt	3b
Lal et al., 2004, USA*	Retrospektive Kohortenstudie Studiengruppe: n=2279 (alle aus Zeitraum) mit invas. z.T: metast. Mammakarzinom Zeitraum: 20 Monate Gewebeprobe: Paraffinschnitte Fixierung: Formalin n.n.def. Test: -Herceptest, - PAtHvysion (dc und sc) Scoring-System: lt Hersteller Verblindeter Vergleich: - Mitgelaufene nach Kit Kontrollen:	Konkordanz IHC und FISHdc: 0 n= 1.006/1.017: 98,9% 1+ n= 693/715: 96,9% 2+ n= 83/313 (13,7%): 26,5% 3+ n= 210/234 (10,3%): 89,7% FISHdc+: 14,3% davon 4,2% als IHC 0 oder 1+ klassifiziert FISHdc und FISHsc Konkordanz: 96%	Sehr gute Fallzahl!	3b
Lebeau et al., 2001, BRD*	Retrospektive Kohortenstudie Studiengruppe: n= 85 konsekutiv mit invasivem Lk+ Mammakarzinom, med. Alter 50J. Zeitraum: 1/1993-11/1997	Technical Failure: Herceptest: 1 FISH aus isolierten Zellkernen: 3 FISH aus Paraffinschnitten: 6		2b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Gewebeproben: Paraffinschnitte Fixierung: neutral gepuffertes Formalin Test: 7 verschied. AK, davon Ergebnisse dargestellt für Cb11(Novocastra), TAB250 (Zymed) und A0485 (DAKO); alle mit AG-Retrieval: Saponin, Digest-All und Microwave; + HercepTest, DAKO PathVysion dc, Visis (auch aus isolierten Nuclei) Scoring-System: n. DAKO und FISH: negativ: </= 4 Signale pro Zellkern Positiv: Ratio>2.0 HA > 10% mit > 10 Signalen pro Zellkern Verblindeter Vergleich: ja Mitgelaufene Kontrollen: -	FISHp+: 27.8% <u>Konkordanz FISHp -IHC:</u> C2+ C3+ T2+T3+ A2+ A3+ H2+ H3+ NA 2 0 1 0 13 0 12 0 LA 0 0 0 0 2 0 2 0 HA 9 11 2 18 2 18 2 18		
Loring et al., 2005, Irland*	Archivmaterial für IHC, prospektiv für FISH und CISH Studiengruppe: n= 119 mit IHC Bestimmung, konsekutiv Zeitraum: 8/2000-4/2003 Gewebeproben: FFPS als TMA Test HercepTest DAKO, FISH = PathVysion Dual Colour CISH = Zymed Scoring-System: a) aber mit 1.8-2.0 als low amplification IHC nach DAKO Verblindeter Vergleich: - Mitgelaufene Kontrollen: -	Konkordanz IHC/FISH ges: 96,3% (106/110) 0/1+ n= 62: 100% (alle negativ) 2+ n= 17: 100% (alle negativ) 3+ n= 31: 87% (27/31)		3b
Lottner et al., 2005*	Retrospektive Kohortenstudie Studiengruppe: n=215, konsekutiv Zeitraum: 1999-2001 Gewebeproben: <i>TMA aus Paraffinschnitten</i>	Konkordanz IHC-FISH: 0/1+ n= 162/168 97% 2+ n= 13/25 52% 3+ n= 22/22 100%		2b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Fixierung: Formalin n.n. def. Test: HercepTest, DAKO und Fluoreszenz -IHC (Ergebnisse n. dargestellt) <i>dc Pathvysion, Abott</i> Scoring-System: <i>IHC n. DAKO, FISH >/=2.0 +,</i> Verblindeter Vergleich: ja Mitgelaufene 4 Zelllinien Kontrollen:			
McCormick et al., 2002, USA *	Retrospektive Kohortenstudie Studiengruppe: n=215 Zeitraum: - Gewebeproben: <i>Paraffinschnitte</i> Fixierung: neutral gepuffertes Formalin Test: HercepTest, DAKO <i>dc Pathvysion, Visis</i> Scoring-System: <i>IHC n. DAKO, FISH >/=2.0 +,</i> <i>>5.0 HA</i> Verblindeter Vergleich: - Mitgelaufene - Kontrollen:	Technical Failure: 17 insgesamt = 198 Konkordanz IHC-FISH: 0 = 96,2% (51/53, 2x FISH LA) 1+= 91,5% (54/59 5x FISH LA) 2+= 42,2% (19/45 14 FISH LA, 26 NA) 3+= 100% (41/41 13 FISH LA, 28 FISH HA) Konkordanz FISH – IHC: FISH+ (33,8%): 89,6% (60/67 7x IHC 0,1+) FISH- (66,2%): 80,2% (105/131, 26xIHC2+)		3b
Mrozkowiak et al., 2004, Polen*	retrospektive Kohortenstudie Studiengruppe: n= 360 Pat. mit invasivemMammakarzinom Zeitraum: - Gewebeproben: <i>Paraffinschnitte</i> Fixierung: 4% gepuffertes Formalin Test: <i>FISH: QBiogene, Oncor, UK</i> IHC: DAKO Herceptest Scoring-System: <i>FISH - </= 4 Signale pro Zellkern und</i> <i>keine Aneuploidie.</i> <i>LA 5-10 Signale, HA>10 Signale</i> <i>IHC: lt DAKO</i> Verblindeter Vergleich: - Mitgelaufene - Kontrollen:	Konkordanz IHC-FISH: 0/1+ (n= 10) 100% 2+ (n= 315) 20% (64x FISH+, 23x Aneuploidie) 3+ (n= 35) 91% (32x FISH+)		3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Owens et al., 2004, USA	retrospektive Kohortenstudie Studiengruppe: n= 6.556 mit FISH und IHC von 116.736 mit FISH gesamt, Referenzlabor Zeitraum: 1999-2002 Gewebeprobe: Paraffinschnitte Fixierung: nicht angegeben Test: -HercepTest mit DAKO Autostainer -Pathvysion Visi Scoring-System: n. DAKO FISH+: >/= 2.0 Ratio Scoring-System: - Verblindeter Vergleich: - Mitgelaufene Kontrollen: It Herceptest	Konkordanz IHC-FISH: 0 (n= 1031/1.075): 95,9% 1+ (n= 1496/1.615): 92,6% 2+ (n= 736/3.159): 23,3% 3+ (n= 648/707): 91,7% Konkordanz FISH+-IHC+ (2+/3+) Gesamt: 89,5% Ratio 2.0-5 (n= 748) 81,7% (62% 2+) Ratio > 5 (n= 799) 96,7% (33% 2+) Mittl. FISH Ratio korreliert pos. mit IHC Score	Empfehlen Überprüfen des Cut-off für FISH und Korrelation mit Outcome	3b
Ogura et al., 2003, Japan*	retrospektive Kohortenstudie Studiengruppe: n= 110 von n = 113 konsekutiven Mammakarzinomen Zeitraum: 6/99-3/2000 Gewebeprobe: Paraffinschnitte Fixierung: 20% Formalin Test: -HercepTest DAKO -PathVysion Vysis Scoring-System: n. Hersteller und FISH + Ratio >/=2.0 Verblindeter Vergleich: - Mitgelaufene Kontrollen: ja	Konkordanz IHC-FISH: 0/1+ (n= 80): 96% 2+ (n= 10): 10% / 90% falls negativ gewertet 3+ (n= 20): 100% Konkordanz FISH-IHC: FISH+: 87,5% FISH-: 89,5% if 2+ als negativ gewertet: FISH+: 83,3% FISH-: 100%		3b
Press et al., 2005, USA	prospektive Kohortenstudie Studiengruppe: n= 2502 mit FISH in 2 zentralen Laboren (Basel und Los Angeles) im Rahmen von 3 BCIRG Studien (005,006,007) von n= 2600 konsekutiven, Vergleich n= 842 mit IHC (HercepTest) und FISH	Konkordanz IHClokal-FISH zentral: 0/1+ (n= 438/457): 95,8% 2+ (n= 21/124): 16,9% 3+ (n= 204/261): 78,2%		2b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Zeitraum: - Gewebeprobe: Paraffinschnitte Fixierung: nicht angegeben Test: zentral: <i>dc PathVision, Abbott-Vysis</i> DAKO HercepTest Scoring-System: FISH: <2 negativ >/= 2 positiv IHC n. DAKO Verblindeter Vergleich: ja Mitgelaufene Kontrollen: ja			
Rossi et al., 2006, Italien	retrospektive Kohortenstudie Studiengruppe: n=81 Zeitraum: 1/2002-1/2004 Gewebeprobe: Paraffinschnitte Fixierung: Formalin n.n.bez. Test: HercepTest PathVysion Scoring-System: n.DAKO, >2.0 FISH+ Verblindeter Vergleich: - Mitgelaufene Kontrollen: ja	Konkordanz IHC-FISH: K0.64 0/1+ (n= 29): 100% 2+ (n= 25): 20% 3+ (n= 21/27): 78% (alle 6 FISH- mit Aneusomie)	Empfehlen alle 2+ und 3+ zu testen	3b
Saez et al., 2006, Spanien	Archivmaterial mit IHC CB11,prospektiv DAKO, FISH und CISH Studiengruppe: n= 200 Zeitraum: 2000-2003 Gewebeprobe: Paraffinschnitte in Formalin fixiert, TMA Fixierung: Test: IHC = CB11 und DAKO; FISH = PathVysion, Vysis, CISH = Zymed Scoring-System: FISH+ > /2.0 Ratio IHC n. DAKO Verblindeter Vergleich: ja, und 2 unabhängige Befunder	Failure FISH: n= 18 Failure HercepT n= 1 Konkordanz IHC-FISH: 0/1+ 2+ 3+ Hercep: 113/115 98% 1/15 7% 39/52 75% CB11: 99/101 98% 0/25 0% 40/56 71,4%	CISH nicht dargestellt	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Mitgelaufene Kontrollen: ja			
Tubbs et al., 2001, USA	retrospektive Kohortenstudie Studiengruppe: n= 145 mit FISH von 400 primären Mammakarzinomen mit IHC insgesamt Zeitraum: - Gewebeproben: <i>Paraffinschnitte</i> Fixierung: <i>Formalin n.n.def.</i> Test: Herceptest, DAKO und CB11 (mit Autostainer) <i>FISH: Probes von Vysis</i> Scoring-System: <i>n. DAKO, FISH + >7= 4.0 Signale pro Zellkern</i> Verblindeter Vergleich: - Mitgelaufene Kontrollen: ja	Konkordanz IHC 2+ und 3+ -FISH: Herceptest: IHC 2+: n= 24 falls -: 87,5%! falls+: 12,5% IHC3+: n= 21/28 75% CB11: IHC2+: n= 17 falls -: 13,7% falls+: 76,4%! IHC3+: n= 17/20 85%		3b
Vocaturio et al., 2006, Italien	Archivmaterial mit Herceptest, prospektiv IHC (CB11), FISH und CISH Studiengruppe: n= 111 , nicht konsekutiv Zeitraum: 1992-2003 Gewebeproben: <i>Paraffinschnitte</i> Fixierung: <i>Formalin n.n.def.</i> Test: IHC = Herceptest DAKO CB11 Novocastra <i>FISH = PathVysion</i> <i>CISH = Zymed</i> Scoring-System: <i>IHC nach DAKO</i> <i>FISH+ Ratio >/=2.0</i> Verblindeter Vergleich: ja Mitgelaufene Kontrollen: ja	Failure: 0% Konkordanz IHC-FISH: Herceptest 0/1+ (n= 53/53): 100%: 2+ (n= 17/36): 47,2% 3+ (n=20/22): 90,9%	CB11 Daten nicht in absoluten Zahlen dargestellt	3b
Dybdal et al., 2005, USA	Bestimmung von FISH aus Archivmaterial mit bereits bestimmtem IHC durch CTA Studiengruppe: n= 623 in 2 Laboren	Technical Failure FISH: n= Konkordanz gesamt IHC-FISH: 82%		3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Zeitraum: - Gewebeprobe: <i>Paraffinschnitte.</i> Fixierung: n. n. def. Test: CTA (4D5 extracellulär, CB11 intracellulär), PathVysion (je Labor mit etwas unterschiedlichem Protokoll) Scoring-System: 0-3 (10%, 1 = inkomplett, 2= komplett schwach, 3= komplett stark) Verblindeter Vergleich: ja Mitgelaufene Kontrollen: +/- Zelllinie	0 (n= 207/214): 97% 1+ (n= 28/30): 93% 2+ (n= 21/88): 24% 3+ (n= 176/197): 89% Reliabilität FISH (Interlabor 2) n= 488 92% K=0.79 37x FISH- in Labor 1 und FISH+ in Labor2: davon 35 IHC3+! LA (1.8-2.2) <2%		
Kakar et al., 2000, USA*	retrospektive Kohortenstudie, Bestimmung aus Archivmaterial von 1 Labor Studiengruppe: n= 117 konsekutive unselektierte Pat. mit inv. Mammakarzinom Zeitraum: - Gewebeprobe: <i>Paraffinschnitte</i> Fixierung: Formalin, n.n.def. Test: CB11, Ventana, dc PathVysion, Vysis Scoring-System: 0, 1+ incomplete, 2+ complete in mind. 10%, 3+ intense complete in >50%. (2+und 3+ als positiv) FISH: + Ratio >/=2.0 Verblindeter Vergleich: ja Mitgelaufene Kontrollen: negative und positive	Technical Failure FISH n= 5 Konkordanz IHC-FISH (n= 112): 0/1+ n= 78/79: 99% 2+ n= 6/17: 35% 3+ n= 14/16: 88% Konkordanz gesamt: 88% <u>Survival –Analyse (v. 65 Pat.) nicht dargestellt</u>		2b
Ridolfi et al., 2000, USA	retrospektive Kohortenstudie Studiengruppe: n= 117 mit FISH von 750 konsekutiven Mammakarzinomen mit IHC Zeitraum: - Gewebeprobe: Paraffinschnitte Fixierung: 10% gepuffertes Formalin Test: A0485 DAKO FISH: <i>Oncor, Ventana und Pathvysion</i> Scoring-System: n. DAKO, FISH+ >4 Signale pro Zellkern oder >/= 2.0 Ratio Verblindeter Vergleich: -	Konkordanz IHC-FISH: 0/1+ (56/57): 98% 2+ (14/38): 37% 3+ (21/21): 100%		3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Mitgelaufene Kontrollen: +			
Tanner et al., 2000, Finnland, Belgien	Prospektive Bestimmung IHC, FISH und CISH Studiengruppe: n=157 , unklar obkonsekutiv Zeitraum: - Gewebeproben: FFT für FISH, Paraffinschnitte für IHC Fixierung: Formalin n.n. def. Test: PathVysion, Visis CB11 , Novocastra Scoring-System: 0-2+ schwache Färbung <50%, 3+ starke Färbung > 50% Fish+ Ratio >/= 2.0 Verblindeter Vergleich: ja Mitgelaufene Kontrollen: -	Konkordanz IHC-FISH: 0/1+ n= 115/126: 91,3% 2+ n= 1/5: 20% (80% bei IHC als negativ gewertet) 3+ n= 25/26: 95%	FISH aus FFT ist nicht Standard, erklärt evtl. die mehr positiven FISH Scoring!?!	3b
Wang et al., 2001, USA	prospektive Kohortenstudie Studiengruppe: n= 151 manuell , 189 ACIS Zeitraum: 8/1998-3/2000 Gewebeproben: Paraffinschnitte Fixierung: 10% neutral gepuffertes Formalin Test: - A0485 DAKO mit AG-Retrieval manuell und mit ACIS - PathVysion Scoring-System: 0 = neg –inkompl. bis zu 10% pos. Zellen 1 + = schwach komplett bis zu 30% pos. Zellen 2+ = mittel komplett 31-50% Zellen 3+ =stark komplett 51-100% Zellen (nur 3+ als positiv gewertet) ACIS: >/= 2.0 als positiv FISH+ Ratio >/= 2.0 Verblindeter Vergleich: -	Konkordanz IHC manuell-FISH: 0 (n= 31): 100% 1+ (n= 45): 98% 2+ (n= 55): 62% 3+ (n= 53/58): 91,4% Konkordanz IHC ACIS-FISH: (n= 115): 92,2% + (n= 4): 89,2% Konkordanz FISH-IHC: Manuell ACIS FISH- (114): 95,6% 93% FISH+ (75): 70,6% 88%		3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Mitgelaufene Kontrollen: ja			
Zhao et al., 2002, USA	Archivmaterial, daraus prospektiv IHC(3 AK), FISH und CISH Studiengruppe: n= 62 Zeitraum: - Gewebeproben: <i>Paraffinschnitte</i> Fixierung: <i>Formalin n.n bez.</i> Test: IHC= Reagenzien von Zymed und DAKO , alle mit AG Retrieval FISH+CISH= Zymed nach Hersteller Scoring-System: IHC n. DAKO FISH >/=2.0 Ratio+ Verblindeter Vergleich: ja Mitgelaufene Kontrollen: ja, positive	Failure: 0% Konkordanz IHC-FISH : 0/1+ 2+ 3+ TAB250: 49/50 98% 4/5 80% 7/7 100% CB11: 47/48 98% 4/7 57% 7/7 100% A0485: 40/40 100% 1/9 11% 11/13 85%	CISH – Ergebnisse nicht dargestellt	

Tabelle 6

Ergebnisse und Evidenzgrad Konkordanz IHC und FISH -II-
nur nach +/- stratifiziert oder nur für IHC2+ oder IHC 3+

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad															
Bartlett et al., 2001, UK	retrospektive Kohortenstudie Studiengruppe: n= 216 Zeitraum: <i>1984-1998</i> Gewebeproben: ‚frozen Sections‘ Fixierung: Formalin n.n. definiert Test: - HercepTest DAKO - CB 11 mit Antigen Retrieval, <i>FISH und Radioimmunhistochemie</i> <i>(Goldstandard)</i> Scoring-System: <i>n. DAKO 2+ und 3+ als positiv gewertet</i> <i>und > 2.0 Ratio Fish +.</i> Verblindeter Vergleich: ja, 2 unabhängige Befunder Mitgelaufene Kontrollen: -	Konkordanz IHC-FISH: <table border="0"> <tr> <td></td> <td>FISH+</td> <td>FISH-</td> </tr> <tr> <td>HercepTest+</td> <td>67%</td> <td>1,7%</td> </tr> <tr> <td>HercepTest-</td> <td>28,6%</td> <td>96%</td> </tr> <tr> <td>CB11+</td> <td>83,3%</td> <td>30,5%</td> </tr> <tr> <td>CB11-</td> <td>14,3%</td> <td>68,4%</td> </tr> </table> <u>Techn. Failure 6xHerceptest,5xCB11, deshalb nicht 100%</u> Reliabilität (Interobserverbeurteilung) Herceptest: 93% CB 11: 62% FISH: 95%		FISH+	FISH-	HercepTest+	67%	1,7%	HercepTest-	28,6%	96%	CB11+	83,3%	30,5%	CB11-	14,3%	68,4%	Radioimmunhisto-chemie als Goldstandard nicht akzeptiert, Ergebnisse nicht dargestellt	3b
	FISH+	FISH-																	
HercepTest+	67%	1,7%																	
HercepTest-	28,6%	96%																	
CB11+	83,3%	30,5%																	
CB11-	14,3%	68,4%																	
Bilous et al., 2003, Australien	Kohortenstudie aus 13 Laboren, zum Teil prospektiv (als Teil der Her2000 Studie) Studiengruppe: n= 1.536 mit Herceptest davon n=109 mit FISH bei IHC 2+ Zeitraum: - Gewebeproben: Paraffinschnitte Fixierung: gepuffertes Formalin n.n.def. Test: HercepTest (manuell oder automatisch) PathVysion Scoring-System: IHC n.DAKO, 2+ und 3+als positiv, FISH+ >2.0 Ratio Verblindeter Vergleich: - Mitgelaufene Kontrollen: ja	Konkordanz IHC2+ -FISH (n= 109 v. n= 206): FISH+ 23% FISH- 77% Interobserver-Reliabilität (14 Pathologen aus 13 Laboren beurteilen 5 IHC-slides): K=0.44		3b															

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Cianciulli et al., 2002, Italien	retrospektive Kohortenstudie	Konkordanz IHC-FISH:		3b
	Studiengruppe:	n= 66		
	Zeitraum:	-		
	Gewebeproben:	Paraffinschnitte	CB11+: FISH+ FISH- 86,2%	
	Fixierung:	neutral gepuffertes Formalin	CB11-: 75,7%	
	Test:	CB11, Biogenex, 300G9 / Ergebnisse nicht dargestellt HercepTest PathVysion	HercepTest +: 100% HercepTest-: 86,4%	
	Scoring-System:	IHC: + if >= 105 der Zellen angefärbt, für HercepTest DAKO System FISH: 2 Ratio +		
	Verblindeter Vergleich:	-		
Mitgelaufene Kontrollen:	SKBR3 Zelllinie als positive Kontrolle und Herceptest Kit Kontrollen +/-			
Couturier et al., 2000, Frankreich	retrospektive Kohortenstudie	Konkordanz IHC-FISH (+/-):		3b
	Studiengruppe:	n=100 mit Mammakarzinom	98/100 = 98%	
	Zeitraum:	-	K= 0.9	
	Gewebeproben:	Paraffinschnitte	12% Überexpression	
	Fixierung:	AFA (75% Äthanol, 5% Essigsäure, 2% Formalin, 18% deionisiertes Wasser)		
	Test:	CB11, ONCOR (Ventana)		
	Scoring-System:	IHC: 0=negativ 1= schwach+ (in >60% der Zellen), 2= stark+ (in >60% der Zellen) FISH > 4 Signale pro Zellekern = +		
	Verblindeter Vergleich:	ja, 2 unabhängige Untersucher		
Mitgelaufene Kontrollen:	normales Epithel als negative Kontrolle		Beachte Scoring! 1+ hat bereits mehr als 60% Zellen +!	
Gancberg et al., 2000, USA	Prospektive Kohortenstudie	Konkordanz IHC-FISH: (2+ als + gewertet)		3b
	Studiengruppe:	n= 160		
	Zeitraum:	-	CB11-: 122/140 Cb11+: 19/20	
	Gewebeproben:	Paraffinschnitt für IHC, frisches Tumorgewebe oder touch preparation für FISH		
Fixierung:	neutral gepuffertes Formalin	TAB250-: 118/130	Werten IHC 2+ einmal positiv und einmal negativ, letzteres hier nicht dargestellt	

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Evidenz- grad
	Test:	TAB250, Zymed, CB11, Novocastra beide mit Antigen Retrieval, polyklonaler Kaninchen Ak (PKAK) FISH: dc Vysis ?	TAB250+: 24/27 PKAK -: 87/93 PKAK+: 29/57		
	Scoring-System:	FISH +: >2.0, IHC: 0, 1+ <= 10% Anfärbung, 2+ = 10-50% Anfärbung, 3+ > 50% Anfärbung			
	Verblindeter Vergleich:	-			
	Mitgelaufene Kontrollen:	positive Zelllinie			
Garcia-Caballero, 2006, Spanien	prospektive Kohortenstudie		Konkordanz IHC2+ (n=60=12%)und	Empfehlen FISH bei 2+ in Referenzlabor	3b
	Studiengruppe:	n= 503 Mammakarzinome aus 10 Laboren, davon 60 mit IHC2+ und FISH.	FISH +: 18% (11/60)		
	Zeitraum:	-			
	Gewebeproben:	Paraffinschnitte			
	Fixierung:	10% gepuffertes Formalin			
	Test:	HercepTest + DAKOCytomation, PathVysion dc			
	Scoring-System:	n DAKO , FISH+: Ratio>2.0			
	Verblindeter Vergleich:	-			
	Mitgelaufene Kontrollen:	+/- vom Testkit, FISH: +inhouse-Kontrolle			
Jacobs et al., 1999, USA	retrospektive Kohortenstudie, Bestimmung aus Archivmaterial von 1 Labor		Technical Failure FISH n= 7 IHC n= 4	Zeit für FISH: 10,4h an 2 Tagen Für IHC: 6h an einem Tag Beurteilung FISH: 8.9min. IHC 2.5 min	2b
	Studiengruppe:	n= 100 konsekutive unselektierte Pat. mit inv. Mammakarzinom	Konkordanz IHC- FISH n=90		
	Zeitraum:	-			
	Gewebeproben:	Paraffinschnitte	IHC+ (n=18/21): 85,7%		
	Fixierung:	Formalin, n.n.def.	IHC- (n= 64/69): 92,7%		
	Test:	- Oncor, INFORM - A0485 mit AG Retrieval			

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Evidenz- grad
	Scoring-System:	0, 1+ schwach, 2+ moderate 3+ stark, circumferentiell, 4+ extrem stark, circumferentiell FISH+ >= 4 Signale pro Zellkern			
	Verblindeter Vergleich:	ja			
	Mitgelaufene Kontrollen:	-			
Harris et al., 2001, USA	retrospektive Kohortenstudie		Technical Failure CB11: n= 2	Auch Survival-Analyse: IHC+ oder FISH+ zeigen beide keine signifikante Korrelation mit OS	3b
	Studiengruppe:	n= 425 von Pat. in Studie mit metast. Mammakarzinom davon n= 207 mit IHC und FISH und ECD (Ergebnisse nicht dargestellt)	Konkordanz IHC-FISH ges.: 77% Konkordanz IHC+-FISH+: 58%		
	Zeitraum:	1988-1995	Konkordanz FISH-IHC-: 87%		
	Gewebeproben:	Paraffinschnitte			
	Fixierung:	Formalin n.n.def.			
	Test:	IHC: CB11 FISH: Inform, Ventana			
	Scoring-System:	IHC n. DAKO, FISH > 4 Signale pro Zelle			
	Verblindeter Vergleich:	-			
	Mitgelaufene Kontrollen:	-			
Lan et al., 2005, Taiwan	prospektive Kohortenstudie		Technical Failure: 5	Aufgrund der Studienergebnisse wird FISH bei IHC2+ bei met. Brustkrebs nun in Taiwan von der Krankenversicherung getragen	3b
	Studiengruppe:	n= 221 von 24 Laboren (1-54 Proben pro Labor) mit IHC2+ und zentral FISH.	Konkordanz IHC2+-FISH±: 44% (96/212)		
	Zeitraum:	3/2003-6/2004	Konkordanz-Rate pro Labor 0%-80%.		
	Gewebeproben:	Paraffinschnitte			
	Fixierung:	nicht bekannt	FISH: 39% LA		
	Test:	dc PathVysion, Visis IHC-Tests verschiedene, n.n.def.			
	Scoring-System:	n. HercepTest, FISH+ >/=2.0 Ratio			
	Verblindeter Vergleich:	-			
	Mitgelaufene Kontrollen:	-			
Larsimont et al., 2002, Belgien	retrospektive Kohortenstudie, Bestimmungen aus Archivmaterial		Konkordanz IHC-FISH:	Empfehlen Einsatz des 'Cocktails', da sensitiver und bei positivem Ergebnis FISH als das kosten-effektivste Modell	3b
	Studiengruppe:	n= 324 Pat. mit invasivem Mammakarzinom in Studie	CB11 -: 86% (43/300 FISH+) CB11+: 88% (21/24 FISH+)		
	Zeitraum:	1989-1996			

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Evidenz- grad
	Gewebeproben:	<i>Paraffinschnitte</i>	TAB250+pAB1-: 94% (15/253 FISH+)		
Fixierung:	nicht bekannt	TAB250+pAB1+: 72% (47/65 FISH+)			
Test:	<i>-CB11 mit Ventana NexEsStaining System, TAB250+pAB1 als Cocktail mit AG-Retrieval mittels Citratpuffer, PathVysion Kit, Vysis</i>	Her2+: CB11: 7%			
Scoring-System:	<i>0, +1 </= 25% pos. Zellen, 2+ </= 50% pos. Zellen, 3+ > 50% pos. Zellen; += 4+ > 75% pos. Zellen (+ = 1-4) FISH+ Ratio > 2.0, > 5.0 HA</i>	TAB250+pAB1: 20% FISH: 20%			
Verblindeter Vergleich:	ja				
Mitgelaufene Kontrollen:	-				
Lopez-Guerrero et al., 2003, Spanien	retrospektive Kohortenstudie		<i>Konkordanz IHC-FISH</i> CB11+ (C): 75% CB11- (A+B): 80,5% HercepTest+ (2+,3+): 73,6% HercepTest- (0,1+): 79,3%	Empfehlen HercepTest und FISH bei 2+	3b
Studiengruppe:	n=102 invasive Mammakarzinome, randomly selected				
Zeitraum:	1982-1993				
Gewebeproben:	<i>Paraffinschnitte</i>				
Fixierung:	Formalin, n.n.def.				
Test:	<i>Cb11, Novocastra und HercepTest FISH: Oncor PCR (Ergebnisse nicht dargestellt)</i>				
Scoring-System:	<i>CB 11 A= kein oder <10% incomplete staining oder > 10% schwache Färbung B=moderate Färbung C= schwache bis starke Färbung in mehr als 10% der Zellen HercepTest nach DAKO FISH: + if Ratio > 2.0 Vergleich n= 104 IHC 3+ aus 67 lokalen Labors (konsekutiv seit Beginn der Studie) mit zentralem Labor im Rahmen der NSABP B-31-Studie</i>				
Verblindeter Vergleich:	-				
Mitgelaufene Kontrollen:	-				

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Paik et al., 2002, USA	retrospektive Kohortenstudie Studiengruppe: n= 104 Zeitraum: 2001 Gewebeprobe: Paraffinschnitte Fixierung: nicht angegeben Test: 1. Lokal HercepTest (DAKO) n= 80 = 77% Andere n= 24 2. zentral Herceptest (DAKO) PathVysion (FISH) Scoring-System: Vom Hersteller Verblindeter Vergleich: - Mitgelaufene Kontrollen: -	Konkordanz Zentrallabor FISH-IHC: 3+ 98/104 = 94,3%		3b
Pauletti et al., 2000, Australien	retrospektive Kohortenstudie Studiengruppe: n=923 konsekutiv Stad. I-III Mammakarzinom Zeitraum: 1987-1991 Gewebeprobe: Paraffinschnitte Fixierung: Formalin, n.n. def. Test: IHC : polyklonaler AK R60 ohne AG- Retrieval FISH: dc PathVysion, Vysis Scoring-System: IHC : 0, 1+ schwach gefärbt, 2+ moderat gefärbt, 3+ stark gefärbt, kein %-Satz gefärbter Zellen FISH: + Ratio > 2.0 oder > 4 Signale pro Zellkern Verblindeter Vergleich: ja Mitgelaufene Kontrollen: ja	Technical Failure: n= 23 (kein Tumorgewebe, oder FISH failure) FISH+ (Ratio> 2.0) 21% IHC +: 17,2% Konkordanz FISH-IHC: FISH+ (n= 178) 57,3% FISH- (n= 678) 93,4%	Korrelation mit Überleben (DFS/OS) FISH erhöht in Bezug auf das Überleben signifikant die prognostische Information von IHC2+oder3+, IHC erhöht die prognostische Information von FISH nur marginal . FISH+/IHC- haben Überleben wie alle FISH+/IHC+ IHC+/FISH- haben Überleben wie IHC-/FISH- PreTest HercepTest und R60 im Vergleich an n= 85 Fällen, finden mehr FP bei HercepTest (25,6%), deshalb R60	2b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Perez et al., 2002, USA	prospektive Kohortenstudie Studiengruppe: n=1575 konsekutiv im Zeitraum mit IHC, davon alle mit 2+ FISH (N=216) Zeitraum: 8/2000-12/2000 Gewebeproben: Paraffinschnitte Fixierung: Test: IHC. HercepTest mit DAKO Autostainer PathVysion dc Scoring-System: DAKO, FISH :? Verblindeter Vergleich: nein Mitgelaufene Kontrollen: Ja	Konkordanz IHC2+-FISH+: 12% (nur high level amplification wird als + gewertet) 25% mit extra Kopien von Her2/neu in Relation zu CEP 17 werden nicht als positiv gewertet	Scoring FISH nur high level als +	3b
Press et al., 2002, USA	retrospektive Kohortenstudie an hinsichtlich der Überexpression und Amplifizierung bekannten Mammakarzinomen Studiengruppe: n= 117, daraus FISH und IHC Zeitraum: - Gewebeproben: Paraffinschnitte Fixierung: Formalin Test: FISH: dc PathVysion, Visis und Ventana Inform IHC R60 polyklonal, 10H8 monoklonal (Ergebnisse nicht dargestellt), DAKO HercepTest, CB11Ventana, Scoring-System: FISH + Ratio > 2.0 oder >4 Signale pro Zellkern Verblindeter Vergleich: - Mitgelaufene Kontrollen: ja	Goldstandard: bekannte Expression/Amplifizierung (wenn Überexpression, auch Amplifizierung n=43, beides negativ n=74) Konkordanz mit FISH PathVysion: K=0.945 Konkordanz mit FISH Ventana: K=0.909 Konkordanz mit CB11: K=0.766 Konkordanz mit HercepTest: K=0.745 (CB11 und HercepTest mit 28% falsch negativen)	R60 und 10H8 mit besserem Kappa. Goldstandard fragwürdig	3b
Roche et al., 2002, USA	retrospektive Kohortenstudie Studiengruppe: n= 119 (erste konsekutive Fälle aus dem Intergroup Trial N9831), Rekrutierung von Pat. mit Her2-neu Überexpression (3+ oder amplifiziert) Zeitraum: Gewebeproben: Paraffinschnitte Fixierung: nicht angegeben	Konkordanz IHC3+-FISH zentral(n=88) 3+ (n=79/88): 89,8%	Aufgrund der Ergebnisse wurde die Aufnahme in die Studie von Bestätigung des Befunds durch ein zentrales Labor abhängig gemacht	3b

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Evidenz- grad
	Test:	<i>lokal</i> IHC: HercepTest 50%, andere 42% FISH: PathVision o. Ventana (ins.7%) Zentral IHC: HercepTest mit DAKO Autostainer/ Fish: PathVision			
	Scoring-System:	IHC: >33% der Zellen stark positiv FISH: > 2.0 oder > 5 Gene Kopies			
	Verblindeter Vergleich:	?			
	Mitgelaufene Kontrollen:	-			
Sauer et al., 2003, Norwegen	retrospektive Kohortenstudie		Technical Failure: n= 13	Survival-Analyse bestätigt die Befunde von Pauletti et al	3b
	Studiengruppe:	n=215 mit IHC, davon 165 mit FISH mit 15 Jahres Follow-up	Konkordanz IHC-FISH:		
	Zeitraum:	1984-1985	0/1+ (120/126) 95,2%		
	Gewebeproben:	Paraffinschnitte			
	Fixierung:	n.angegeben	2+/3+ (21/26) 80,7%		
	Test:	CB 11, Biogenex INFORM, Ventana			
	Scoring-System:	n. DAKO und FISH+ > 4 Signale pro Zellkern			
	Verblindeter Vergleich: Mitgelaufene Kontrollen:	- ja			
Seidman et al., 2001, USA	prospektiv, Phase II Studie		Herceptest zeigt im Vergleich zu CB11 9 mehr 2+/3+:	2+ /3+ nicht in absoluten Zahlen Ergebnisse für Pab1 und Tab250 nicht mit FISH korreliert	2b
	Studiengruppe:	alle n= 95 Pat. mit metast. Mammakarzinom und bis zu 3 vorherigen Chemotherapien, Behandlung mit Trastuzumab und Paclitaxel mit und ohne Her2- Überexpression	Konkordanz IHC- FISH+:		
	Zeitraum:	-	Cb11 0/1+: 36/42 85,7%		
	Outcome-Parameter:	Overall Response	CB11 2+/3+: 34/36 94,4%		
	Gewebeproben:	Paraffinschnitte	HercepT. 0/1+: 30/33 90,9%		
	Fixierung:	Formalin, nicht näher angegeben	HercepT. 2+/3+: 37/45 82,2%		
	Test:	DAKO HercepTest,Pab 1,TAB250,Cb11 FISH: dc PathVysion			
	Scoring-System:	nach DAKO FISH: >/= 2 Ratio ist positiv			

Autor/ Jahr/Land	Studienbeschreibung		Ergebnisse	Bemerkungen	Evidenz- grad
	Verblindeter Vergleich:	ja und Beurteilung durch 2 Pathologen			
	Mitgelaufene Kontrollen:	-			
Tawfik et al., 2006, USA	retrospektive Kohortenstudie		Konkordanz FISH-IHC:		3b
	Studiengruppe:	n= 247			
	Zeitraum:	6/1998-3/2004	FISH+: (39/47) 83%		
	Gewebeproben:	Paraffinschnitte	FISH-: (193/200) 96,5%		
	Fixierung:	10% neutral gepuffertes Formalin			
	Test:	<i>HercepTest DAKOCytomation mit ACIS Automation System, ChromaVision PathVysion, Vysis</i>	Diskordanz bei n= 15, davon 6 mit FISH und IHC borderline +. 12 v 15 waren Core Biopsien.		
	Scoring-System:	FISH + >/=2.0, ACIS + >/=2.2			
	Verblindeter Vergleich:	-			
	Mitgelaufene Kontrollen:	-			

5.3.1 Ausgeschlossene Studien zur Konkordanz IHC und FISH

Tabelle 7

Ausgeschlossene Studien zur Konkordanz IHC-FISH

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Arnould et al., 2003, Frankreich	Vergleich CISH-FISH-IHC	Nur 47 mit Formalin fixiert, FISH zum Großteil aus FFT
Bankfalvi et al., 2000, BRD	Vergleich Her2-Bestimmung mit: HercepTest (DAKO) -A0485 (automatisch und manuell) -CB11 -Oncor, Ventana -cdPCR und CGH	Fallzahl < 50 (40)
Bilous et al., 2006, Australien	Vergleich IHC,FISH, CISH	Fallzahl < 50 (49)
Bussolati et al., 2005, Italien	Neue Her-2 Detektion mit Biotin-labeled Trastuzumab	Kein Vergleich IHC-FISH
Choi et al., 2003, USA/Korea	Vergleich von IHC und FISH bei Frauen aus USA und Koreanerinnen	Fokussiert auf Unterschiede der Kohortengruppen
Dagrada et al., 2003, Italien	Vergleich verschiedener AK und FISH vor und nach Chemotherapie	Fokussiert auf Pretreatment-Posttreatment-Vergleich
Dandachi et al., 2004, Österreich,	Vergleich IHC; FISH, CISH	Fallzahl < 50 (40)
Diaz et al., 2004, USA	Vergleich IHC, FISH, CISH	Fallzahl < 50
Di Palma, 2007, UK	s.o.	Fallzahl < 50
Gancberg, 2002, Belgien	Vergleich FISH, CISH, IHC	Stark abweichendes Scoring IHC
Ginestier et al., 2004 Frankreich	Vergleich IHC -FISH	TMA aus FFP (nicht üblich) Fixierung nicht angegeben
Gouvea et al., 2006, Brasilien	Vergleich FISH mit: CB11 Biogenex und Novocastra (beide ohne Antigenretrieval) 4D5 Genentech (Trypsin) A0485 (Steamer) HercepTestKIT	Fallzahl < 50 (36)
Harbeck et al., 1999, BRD	Vergleich IHC Ab-3 AK und FISH	Kein gebräuchlicher AK
Hatanaka et al., 2001, Japan	Vergleich IHC manuell-automatisch	Kein Vergleich FISH
Jacobs et al., 1999, USA	Vergleich IHC - FISH	Fallzahl kleiner 50 (48)
Jimenez et al., 2000, USA	Vergleich IHC- FISH	Fallzahl < 50 (n= 34)
Kakar et al., 2000, USA	Vergleich FISH PAtHvysion und INFORM	Kein Vergleich IHC
Kobayashi et al., 2002, Japan	Vergleich IHC und FISH	Kommerzieller japanischer AK, hier nicht gebräuchlich
Kunitomo et al., 2002, Japan	Vergleich IHC und FISH	Kommerzieller Japanischer AK, hier nicht gebräuchlich
Li-Ning et al., 2005, USA	Vergleich IHC, FISH und CISH	Fallzahl < 50
Merkelbach-Bruse et al., 2003, BRD	Vergleich IHC, FISH und PCR	Fokus auf PCR (Zahlenangaben)
O'Malley. et al., 2001, USA	Vergleich IHC und PCR	Kein Vergleich IHC-FISH
Peiro et al., 2007, Spanien	Vergleich IHC,FISH und CISH	Fallzahl FISH< 50

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Perez et al., 2006, USA	Interlaborvergleich FISH, IHC	Absolute Zahlen für Konkordanz IHC-FISH fehlen
Sapino et al., 2003, Italien	Vergleich IHC-CISH	Kein Vergleich IHC-FISH
Selvarajan et al., 2004, Singapore	Vergleich HercepTEst und andere IHC	Kein Vergleich IHC-FISH
Sidoni et al., 2006, Italien	Vergleich HercepTest und FISH bei vorher zweifelhaften CB11 Ergebnissen bei hoher Hintergrundfärbung	Fallzahl < 50 (technisches Versagen FISH 26% von n=65), spezielle Situation bei zweifelhaftem CB11
Tapia et al., 2004, Schweiz	Vergleich FISH-IHC an TMA	Von n= 2.197 FISH bei 1.631, in absoluten Zahlen den IHC-Scores nicht zuordenbar!
Thomson et al., 2001, Kanada	Vergleich IHC 4 AK und FISH	Interobserververgleich, z.T. kein Konsensus, deshalb keine absoluten Zahlen
Tsuda et al., 2001, Japan	Vergleich IHC und Southern Blot Hybridisierung	Kein Vergleich IHC-FISH
Tubbs et al., 2002, USA	Vergleich FISH – GoldEnhance Bright-Field-Methode	Kein Vergleich FISH-IHC
Wixom et al., 2004, USA	Vergleich FISH, CISH, IHC	Fallzahl <50 (n= 33)

5.4. Konkordanz von FISH und CISH

5.4.1 Ergebnisse und Evidenzgrad FISH-CISH

Tabelle 8

Ergebnisse und Evidenzgrad FISH-CISH

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Arnould et al., 2003, Frankreich	Archivmaterial mit FISH und IHC, daraus prospektiv CISH	Failure: CISH: 5,1% (4/79)	62x Centromer 17 bestimmt, davon Polysomie in 62,5% (10/16) bei IHC 2+ Tumoren Polysomie nur in einem diskordanten FISH/CISH-Fall ursächlich vermutet Unterschiedliche Gewebebefixierung	3b
	Studiengruppe: n= 79 aus 8 Laboren	alle aus Gewebe mit Alkohol-Formalin- Essigsäure-Fixierung		
	Zeitraum: -	Sensitivität CISH: 97% Spezifität CISH: 95%		
	Gewebeproben: IHC und CISH aus FFPS (47 neutral gepuffert, 10 Hollands Bouin, 22 Alkohol-Formalin-Essigsäure) FISH 65x aus FFT und 14x FFPS	Sensitivität CISH bei IHC 0/1+/3+: 95% Spezifität CISH bei IHC 0/1+/3+: 100%		
	Test: FISH = PathVysion oder Ventana CISH = Zymed IHC = DAKO 54x CB11 25x 10x? Alle nach Hersteller- Protokoll ausgeführt	Sensitivität CISH bei IHC 2+: 100% Spezifität CISH bei IHC 2+: 85%		
	Scoring-System: FISH= a) 62x b) 14x CISH= c) und HAC auch große Signalcluster IHC = d)	Konkordanz FISH-CISH: a) alle + (LA/HA) und alle - 96% (72/75) K=0.97 b) differenziert nach NA, LA, HA 93,3% (70/75) K=0.88		
Verblindeter Vergleich: ja	<u>Konkordanz Gruppe 1 IHC (0/1+, n=27):</u> 92,6% (25/27)			

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Mitgelaufene Kontrollen: nein	<p>alle neg., 2x CISH ohne Signal</p> <p><u>Konkordanz Gruppe 2 IHC (2+,n=29):</u> 82,8% (24/29) FISH 14x neg, CISH 12x neg, CISH 2x HA bei LA FISH, 1x CISH ohne Signal</p> <p><u>Konkordanz Gruppe 3 IHC (3+,n=23):</u> 91,3% (91/93) 1xFISH HA, CISH NA, 1xCISH ohne Signal</p>		
Bhargava et al., 2005, USA	<p>Archivmaterial mit FISH, daraus prospektiv CISH</p> <p>Studiengruppe: n= 113 von 2.000 Fällen, die für FISH getestet wurden, v.a. mit low-level- Amplifikation, 1 Labor</p> <p>Zeitraum: -</p> <p>Gewebeproben: FISH und CISH aus FFPS in TMA (0.6mm) jeweils 4 Proben pro Fall</p> <p>Test: FISH=PathVysion Sonde dc, DAPI (homebrew?) CISH=Zymed mit Chr. 17 nach Herstellerangabe jeweils mind. 30 Zellkerne beurteilt</p> <p>Scoring-System: FISH= a) LA und HA als + zusammengefasst CISH= d) Diff. nach Ratio eigenes System</p> <p>Verblindeter Vergleich: -</p>	<p><u>Failure:</u> FISH: 7% (8/113) CISH: 5,3% (6/113) 3x beide</p> <p><i>Sensitivität CISH (+/-):</i> 100%</p> <p><u>Spezifität CISH (+/-):</u> 100%</p> <p><u>Konkordanz ges. +/-:</u> 100%</p> <p>K=1</p> <p>63 > 2 = amplifiziert 39 < 2 = non-ampl.</p> <p><u>Konkordanz diff. nach Ratio:</u> NA <2 (n=39): 100%</p>	<p>Studie fokussiert auf low-level Amplifikation</p> <p>Eigene Definition von Borderline und Low-Level sowie High-level Ratio</p> <p>Gründe für Failure: Verlust der Microarray-Cores, weniger als 30 Zellkerne vorhanden kein Her2 oder Chr. 17 Signal</p>	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Mitgelaufene Kontrollen: nein	Borderline 2.0-2.5 (n=6): 50% LA 2.6-3.9 (n=16)= 38% HA >/=4 (n=41)= 88% Upgrade CISH:11 Downgrade CISH: 6		
Bilous et al., 2006, Australien	Archivmaterial mit FISH und IHC 2+ und 3+, prospektiv CISH Studiengruppe: n= 49 aus 5 Laboren, CISH jeweils in 2 Laboren bestimmt Zeitraum: - Gewebeprobe: FFPS Test: IHC = DAKO FISH = PathVysion dc CISH = Zymed alle nach Herstellerangaben Scoring-System: a) c) d), jeweils 30 Zell-Kerne beurteilt Verblindeter Vergleich: ja Mitgelaufene Kontrollen: -	Failure: CISH 2% (1/50) Konkordanz FISH-CISH: a) alle 83,7% (41/49) b) alle amplifizierten 90% (27/30) c) alle nicht-amplifizierten 86% (19/22) <u>die größte Diskordanz im LA-Bereich :</u> 60%(n= 6/10) Upgrade CISH: 3 Downgrade CISH: 7 <u>Inter-Laborvergleich CISH:</u> Reliabilität ges: 79,6% (39/49)	Keine Chr. 17 Best. für CISH	3b
Dandachi et al., 2004, Österreich	Archivmaterial, prospektiv für IHC und CISH, z.T. FISH Studiengruppe: n= 171 IHC und CISH, davon 40 mit FISH Zeitraum: 1993-1999 Gewebeprobe: FFPS Test: IHC = DAKO, FISH = PathVysion dual colour beide nach Hersteller CISH = Zymed DNA Probe und weitere Reagenzien (Böhringer, DAKO)	Failure: CISH=0% FISH=5%(2/40) Sensitivität CISH: 100% Spezifität CISH: 100% Konkordanz FISH-CISH: 100% K=1	Sehr kleines n für FISH	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Scoring-System: IHC = e) FISH = b) mind.60 Zellen CISH = c) Verblindeter Vergleich: ja Mitgelaufene Kontrollen: -	<u>Konkordanz IHC 3+ mit FISH/CISH:</u> 92% (23/25) 2x FISH/CISH negativ		
Diaz et al., 2004, USA	Archivmaterial, prospektiv für IHC und CISH, z.T. FISH Studiengruppe: n= 41 Zeitraum: 2000-2001 Gewebeprobe: FFPS nach Alkohol-Formalin-Puffer in TMA (1,5mm), jeweils 2 Proben Test: IHC = A0485 AK v. DAKO mit eigenem Protokoll, FISH = PathVysion dc CISH = Zymed beide nach Hersteller Scoring-System: IHC = e) FISH = b) mind.60 Zellen LA und HA zusammengefasst CISH = c) Verblindeter Vergleich: ja, in 2 Laboren Mitgelaufene Kontrollen: -	<u>Failure:</u> CISH = 22% (9/41) FISH = 12,2% (5/41) IHCe = 7,3% (3/41) <u>Konkordanz FISH-CISH:</u> 87,5% (37/41) 4x FISH- (Polysomie) und CISH LA + Interlaborkonkordanz: 100%	Sehr kleines n	3b
Di Palma et al., 2007, UK	Archivmaterial mit IHC und z.T. FISH, prospektiv für CISH Studiengruppe: n=161 IHC, davon n= 24 mit FISH Zeitraum: - Gewebeprobe: FFPS Test: IHC = DAKO, FISH = PathVysion dual colour CISH = Zymed, Invitrogen Scoring-System: jeweils das vom Hersteller b)c) CISH: Amplifikation falls >50% der Zellkerne mit Signalklustern oder >5 Signalen Keine Amplifikation falls >50% der Zellkerne mit max.1-2 kleinen Signalen Verblindeter Vergleich: nein, aber CISH wird von 2 Pathologen unabhängig voneinander beurteilt Mitgelaufene Kontrollen: -	<u>Failure:</u> 0% <u>Sensitivität CISH:</u> 100% <u>Spezifität CISH:</u> 100% <u>Konkordanz FISH-CISH:</u> 100% <u>Konkordanz ges. IHC-CISH:</u> 85,6% (141/161) Diskordant 3+ IHC und – CISH (n= 7) alle durch FISH bestätigt. Diskordant – IHC und + CISH (n= 1) durch FISH bestätigt	Sehr kleines n	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Gong et al., 2005, USA	Archivmaterial mit FISH, prospektiv für CISH Studiengruppe: n= 80 konsekutive Fälle Zeitraum: - Gewebeprobe: FFPS Test: FISH= PathVysion dual colour, Visis; CISH = Zymed Scoring-System: FISH = a) mind. 60 Zellen beurteilt (Polysomie Chr. 17: > 6 Signale) CISH = c) Verblindeter Vergleich: nein, aber CISH wird unabhängig von 3 Pathologen beurteilt Mitgelaufene Kontrollen: ja, FISH+/-	Failure: 0% Konkordanz FISH-CISH alle: a) +/-: 91% (73/80) b)diff. In low-level und high-level Amplifikation: 83,8% (67/80) Interobserver-Reliabilität CISH (+/-): 91% (73/80) Konkordanz FISH-CISH pro Observer: 94%, 95% und 96% Diskordanz Interobserver: 6x low-level/high level (minor) 7x pos.-neg.(major)		3b
Gupta et al., 2003, USA	Prospektive Bestimmung FISH/CISH Studiengruppe: n= 31 von ca. 50 invasiven Tumoren im Zeitraum mit CISH-Bestimmung, alle amplifizierten eingeschlossen Zeitraum: 5 Monate Gewebeprobe: FFPS Test: FISH= PathVysion, Visis, dc mind. 60 Zellen beurteilt CISH= Zymed, mind. 50 Zellen beurteilt Scoring-System: FISH = a) CISH = c) Verblindeter Vergleich: ja; CISH wird unabhängig von 2 Pathologen beurteilt Mitgelaufene Kontrollen: FISH: positive Kontrollen von Visis CISH: invasiver Tumor Her-2/neu-IHC und FISH+, normales Brustgewebe	Failure: 0% Sensitivität CISH: Spezifität CISH: Konkordanz FISH-CISH: 83,9% (26/31) Upgrade CISH: 5 davon 4 mit lowlevel amplification wegen Polysomie Chr.17 1 CISH falsch positive bei high background-Signal – bei Wiederholung negativ Downgrade CISH: 0		3b
Hanna et al., 2006,	Prospektive Bestimmung IHC, FISH/CISH Studiengruppe: n= 254 , unklar, ob konsekutiv	Failure: 0%	CISH nicht auf Polysomie 17 geprüft	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
USA	Zeitraum: - Gewebeprobe: FFPS Test: IHC = Herceptest, DAKO, FISH = PathVysion dual colour nach Angaben der Hersteller CISH = local protocol und Zymed- detection Kit Scoring-System: IHC = e) FISH = a) CISH = <6 Signale = NA, >6 = ampl. Verblindeter Vergleich: ja Mitgelaufene Kontrollen: -	Konkordanz FISH-CISH ges: 95,3% (242/252) K= 0.91 Konkordanz in Gr. 1 (IHC 0/1+; (n=69): 97% (2/69) 2x – FISH und low-level CISH+ Konkordanz in Gruppe 2 (IHC 2+; n=135): 93% (126/135) 8x FISH+ und CISH- 1x FISH- und CISH+ Konkordanz in Gruppe 3 (IHC 3+; (n=50): 98% (49/50) 1x FISH- und CISH+		
Isola, Tanner et al., 2004	Archivmaterial für jeweils die Hälfte von FISH/CISH Studiengruppe: n ges = 197, n= 98 von Finnland mit CISH (FISH für Studie), n= 99 von UK mit FISH (CISH für Studie), jeweils 50% amplifizierte für das bereits durchgeführte Verfahren Zeitraum: - Gewebeprobe: FFPS Test: FISH= PathVysion detection Kit, pretreatment mit VP2000 Prozessor CISH= Zymed Reagenzien und Roche/Power-vision detection Kit Scoring-System: FISH = a), CISH = c), bei Borderline-Kopie-Zahl, Chr. 17 Bestimmung Verblindeter Vergleich: ja Mitgelaufene Kontrollen: bei CISH: FISH+/-	Failure: CISH: 3% (3/99) FISH: 2% (2/98) Konkordanz FISH/CISH: 93,8% (180/192) K= 0.875 92x beide negativ, 88x beide positiv Upgrade CISH: 8 7x fehl. Chr. 17 Info CISH Downgrade CISH: 4 2x CISH falsch negativ wegen zu wenigen Tumorzellen im Präparat Mediane Anzahl von Genkopien: FISH(3.5-14.8) CISH(4.3-14.2) K= 0.75 (p<0.0001) Interlabor-Reliabilität CISH: 95,9% (186/194)	CISH für Chr. 17 Centromer in 7 Fällen	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Laakso et al., 2006, Finnland	Prospektive Bestimmung Studiengruppe: n= 40 , CISHdc-Routinediagnostik und 4 Zell-Linien Zeitraum: - Gewebeprobe: FFPS Test: FISH = Kommerzielle Sonden (Vysis) CISH = PBS, Sonde: I Invitrogen/Roche, eigenes Protokoll Dc! Scoring-System: FISH = a) mind. 20 Zellen CISH = d) mind. 100 Zellen beurteilt Verblindeter Vergleich: ja Mitgelaufene Kontrollen: -	Failure:0% Konkordanz FISH/CISH: 91% (40/44) K=0.82	Sehr kleines n	3b
Li-Ning et al., 2005, USA	Archivmaterial mit IHC Studiengruppe: n= 55 , per Zufall ausgewählte Tumore, davon n=32 mit FISH und CISH (aus 2 Laboren) Zeitraum: - Gewebeprobe: FFPS Test: FISH= PathVysion dc CISH= Zymed Scoring-System: a)c) Verblindeter Vergleich: ja Mitgelaufene Kontrollen: -	Failure: n.ber. Konkordanz FISH/CISH: 96,9% (31/32) 1xFISH negativ und CISH + low-level mit Disomie Chr. 17	Sehr kleines n	3b
Loring et al., 2005, Irland	Archivmaterial für IHC, prospektiv für FISH und CISH Studiengruppe: n= 119 mit IHC Bestimmung, konsekutiv Zeitraum: 8/2000-4/2003 Gewebeprobe: FFPS als TMA Test: IHC = DAKO, FISH = PathVysion dual colour CISH = Zymed Scoring-System: a) aber mit 1.8-2.0 als low amplification Verblindeter Vergleich: -	CISH/FISH-Failure: 7,6% (9/119) Konkordanz CISH/FISH ges: 99% (109/110) Konkordanz CISH/FISH in Gruppe1 IHC (0/1+,n=62): 100% (alle negativ) Konkordanz	Korrektur CISH Centromer 17 ?	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Mitgelaufene Kontrollen: -	CISH/FISH in Gruppe 2 IHC (2+, n=17): 100% (alle negativ) <u>Konkordanz</u> CISH/FISH in Gruppe 3 IHC (3+, n= 31): 96,8% (30/31) 4xFISH- und 3x CISH negativ, DAKO-Test auch bei Wiederholung3+		
Park et al., 2003, Korea	prospektiv für FISH und CISH aus Archivmaterial Studiengruppe: n= 188 Zeitraum: - Gewebeproben: FFPS in TMA Test: FISH = Pathvysion dual colour CISH = Zymed Scoring-System: FISH = >=2xHer2/neu Signale als Centr. 17 Signale = ampl. CISH = >4 Signale oder große Signalcluster in mehr als 50% der Zellkerne Verblindeter Vergleich: - Mitgelaufene Kontrollen: -	Failure: 0% <u>Konkordanz CISH/FISH ges:</u> 94,1% (177/188) K= 0.838 7xFISH+ und CISH- 4xFISH- und CISH+		3b
Peiro et al., 2007, Spanien	Archivmaterial mit IHC Studiengruppe: n= 40 mit FISH und CISH Zeitraum: - Gewebeproben: FFPS Test: FISH = Dako Cytomation dual colour, CISH = Zymed Scoring-System: FISH = a) für mind. 100 Zellkerne, CISH = c) für mind. 200 Zellkerne Verblindeter Vergleich: - Mitgelaufene Kontrollen: -	Failure: CISH 1,9% (3/162) FISH 0% <u>Konkordanz CISH/FISH:</u> 94,6% (35/40) K=0.841 2xFISH+ und CISH- 3xFISH- und CISH+		3b
Saez et al., 2006, Spanien	Archivmaterial mit IHC CB11,prospektiv DAKO, FISH und CISH Studiengruppe: n= 200 Zeitraum: 2000-2003 Gewebeproben: FFPS in TMA	Failure: CISH: 10% (20/200) FISH: 9% (18/200)	Keine Chromosom17 Information für CISH 2x Miss-Interpretation	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Test: IHC = CB11 und DAKO; FISH = PathVysion, Vysis, CISH = Zymed Scoring-System: FISH = a) für mind. 60 Zellkerne, CISH = c) Verblindeter Vergleich: ja, und 2 unabhängige Befunder Mitgelaufene Kontrollen: -	Sensitivität CISH: 97,5% Spezifizität CISH: 94% Konkordanz CISH/FISH: 94,8% (165/174) K=0.86 (39+/126-) 1xFISH+ und CISH- 8xFISH-und CISH+ (7xFISH Polysomie) <u>Interobserver- Reliabilität CISH:</u> 97,5% (195/200)	Ergebnisse für IHC- Konkordanz nicht dargestellt	
Tanner et al., 2000, Finnland, Belgien	Prospektive Bestimmung FISH und CISH Studiengruppe: n= 157, unklar ob konsekutiv Zeitraum: - Gewebeprobe: FFT für FISH, FFPS für CISH Test: FISH = verschied. Reagenzien u.a. Vysis CISH = Zymed IHC = CB11 Scoring-System: FISH = a) für mind. 250 Zellkerne CISH = c) IHC = e) Verblindeter Vergleich: ja Mitgelaufene Kontrollen: -	<u>Failure:</u> CISH: 2% (3/160) Konkordanz CISH/FISH ges: 93,6% (147/157) K=0.81 10x FISH+ und CISH- <u>Konkordanz IHC Gruppe1 (0/1+):</u> FISH-IHC: 91,3% (115/126) CISH-IHC:98,4% (124/126) <u>Konkordanz IHC Gruppe2 (2+):</u> FISH-IHC:80% (4/5) CISH-IHC:100% (5/5) CISH alle negativ <u>Konkordanz IHC Gruppe3 (3+):</u> FISH/CISH-IHC:96% (25/26) 1x FISH und CISH neg.	FISH aus FFT ist nicht Standard, erklärt evtl. die mehr positiven FISH; 9 von 10 positiven FISH bei negativem CISH hatten low- level/Borderline Ampl. keine Information über Centromer 17 CISH	3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Vocaturio et al., 2006, Italien	Archivmaterial mit IHC (DAKO), prospektiv IHC (CB11), FISH und CISH Studiengruppe: n= 157 Zeitraum: 1992-2003 Gewebeproben: FFPS Test: IHC= DAKO nach Hersteller CB11 mit eigenen Reagenzien FISH= PathVysion CISH= Zymed beide nach Hersteller Scoring-System: FISH=a) CISH=c) zusätzlich Chr. 17 Best. für LA IHC= e) Verblindeter Vergleich: ja und CISH unabhängig von 2 Befundern Mitgelaufene Kontrollen: ja	Failure: 0% Konkordanz CISH/FISH ges: 94,6% (105/111) K=0.88 6x FISH+ und CISH- Konkordanz IHC Gruppe1 (0/1+): FISH/CISH=100% alle negativ Konkordanz IHC Gruppe2 (2+): FISH/CISH:100% 17 ampl., 19 nicht amplifiziert Konkordanz IHC Gruppe3 (3+): FISH/CISH: 72,7% (16/22) 6x FISH+ und CISH- Interobserver-Reliabilität CISH: K=0.96		3b
Wixom et al., 2004, USA	Archivmaterial, daraus prospektiv IHC, FISH und CISH Studiengruppe: n= 33 Zeitraum: - Gewebeproben: FFPS Test: IHC = Reagenzien von Zymed (auch AK) und DAKO FISH = PathVysion n Hersteller CISH = Reagenzien ymed und PBS- Waschung Scoring-System: FISH = a) LA und HA als + CISH = c) IHC = e) Verblindeter Vergleich: ja Mitgelaufene Kontrollen: ja	Failure: 0% Konkordanz CISH/FISH ges: 96,8% (30/31) K=0.88 Amplifikation: 1x FISH- und CISH+ low level Konkordanz CISH-IHC bei 0, 2+ und 3+: 100% Konkordanz CISH-IHC Gruppe bei 1+: 55,6% (5/9) 4x CISH pos.	Sehr kleines n	3b
Zhao et al.,	Archivmaterial, daraus prospektiv IHC(3 AK), FISH und CISH	Failure: 0%		3b

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
2002, USA	Studiengruppe: n= 62	Konkordanz CISH/FISH ges: 100% (30/31) K= 1 Amplifikation: 19% (12/62) Konkordanz IHC-AK-CISH: a) TAB250: 97% b) CB11: 94% c) A0485: 84%		
	Zeitraum: -			
	Gewebeproben: FFPS			
	Test: IHC = Reagenzien von Zymed (auch AK) und DAKO FISH+CISH = Zymed nach Hersteller			
	Scoring-System: FISH =a) CISH = c) und Chr. 17 Best. bei 3-5 Signalen IHC = e)			
	Verblindeter Vergleich: ja			
Mitgelaufene Kontrollen: ja, positive				

5.5 Her-2/neu-Bestimmung im Vergleich Stanzbiopsie-definitives Operationspräparat

5.5.1 Ergebnisse und Evidenzgrad Her-2/neu-Bestimmung im Vergleich Stanzbiopsie – definitives Operationspräparat

Tabelle 9

Her-2/neu Bestimmung im Vergleich Stanzbiopsie-definitives Operationspräparat

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
Burge et al., 2006, USA	Erhebung retrospektiv	Konkordanz ges.: IHC: 96% (78/81) FISH: 100% Konkordant +: IHC (3+): 100% (4/4) Konkordant -:	FISH aus Stanze nur in 22% (n= 18) vorhanden Anzahl der Stanzzyylinder nicht angegeben	2b
	Studiengruppe: n= 81 , alle mit CNB und Exzisionsbiopsie im Haus, Pat. mit neoadj. Therapie ausgeschlossen			
	Zeitraum: 10/2002-2/2004			
	Biopsieart: 14g CNB, Anzahl Zylinder nicht angegeben			
Test: HercepTest (DAKO) und Fish (n= 18, PathVysion)				

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad												
	Verblindeter Vergleich: ja	IHC: 96% (74/77) 3 negativ in Biopsie und positiv in FSE														
Cahill et al., 2006, Irland	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 20%;">Erhebung</td> <td>prospektiv</td> </tr> <tr> <td>Studiengruppe:</td> <td>n= 95, konsekutiv, alle Tumorstadien</td> </tr> <tr> <td>Zeitraum:</td> <td>3 Jahre</td> </tr> <tr> <td>Biopsieart</td> <td>18g CNB, US-gest., 3-5 Stanzzyylinder</td> </tr> <tr> <td>Test:</td> <td>IHC, n.n spez.</td> </tr> <tr> <td>Verblindeter Vergleich:</td> <td>nicht angegeben</td> </tr> </table>	Erhebung	prospektiv	Studiengruppe:	n= 95 , konsekutiv, alle Tumorstadien	Zeitraum:	3 Jahre	Biopsieart	18g CNB, US-gest., 3-5 Stanzzyylinder	Test:	IHC, n.n spez.	Verblindeter Vergleich:	nicht angegeben	<p><u>Konkordanz ges.:</u> IHC: a) 100% (+/-) b) 64% (diff. nach 0 bis 3+)</p> <p>Upgrade CNB: 27% Downgr. CNB: 7% (gegenüber FSE)</p> <p><u>Konkordant +:</u> IHC (3+): 87,5% (16/14) 2 positiv in CNB, aber nicht in FSE</p> <p><u>Konkordant -:</u> <u>IHC (0):</u> 75% (20/15) 5 negativ in CNB, aber 1+ oder 2+ in FSE</p> <p><u>IHC (1+):</u> 66% (41/27) 13x 1+ in CNB, aber 10x negativ in FSE 3x 2+ in FSE</p> <p><u>IHC (2+):</u> 28% (18/5) 13x 2+ in CNB, aber 7x negativ in FSE 6x 1+ in FSE</p>	CNB hat Tendenz zu höherem Scoring+	3b
Erhebung	prospektiv															
Studiengruppe:	n= 95 , konsekutiv, alle Tumorstadien															
Zeitraum:	3 Jahre															
Biopsieart	18g CNB, US-gest., 3-5 Stanzzyylinder															
Test:	IHC, n.n spez.															
Verblindeter Vergleich:	nicht angegeben															

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad	
Cavaliere et al., 2005, Italien	Erhebung	prospektiv	<u>Konkordanz ges:</u> IHC: 89,7% (61/68) T1: 92,9% (39/42) T2: 84,61% (22/26)	Keine Differenzierung von 0-3+; Cut-off for IHCe Positive: =>10%. Keine Aussage zu Upgrade oder Downgrade möglich Stratifizierung nach T1 und T2	2b
	Studiengruppe:	n= 68 , konsekutiv, T1 und T2-Tumoren Pat. mit neoadj. Therapie ausgeschlossen			
	Zeitraum:	-			
	Biopsieart	18g CNB, 2-6 Stanzzyylinder			
	Test:	IHC			
Verblindeter Vergleich:	ja				
Connor et al., 2002, USA	Erhebung	-	<u>Konkordanz ges:</u> <u>IHC:</u> 90,9% (40/44) <u>Diskonkordant +:</u> 2 <u>Diskordant -:</u> 2	Cut-off für IHCe Positive >2%???	3b
	Studiengruppe:	n= 44 , keine Pat. mit neoadj. Therapie, T1-T3 Tumoren			
	Zeitraum:	-			
	Biopsieart	14g CNB, US-gest. oder 11g VAB handheld stereotactic			
	Test:	IHC			
Verblindeter Vergleich:	nicht an gegeben				
Erggelet et al., 2007, BRD	Erhebung	retrospektiv	<u>Konkordanz gesamt::</u> 46% IHC 0/1+: 73,6% IHC 2+: 33,3% IHC 3+: 85% Anzahl 3+: CNB 20/18% FSE 31 (28%) Anzahl 2+: 18 18 Anzahl 1+: 50 (46%) 24 (22%) Anzahl 0: 22 37	CNB hat Tendenz zu niedrigerem Scoring	3b
	Studiengruppe:	n= 110 , alle mit CNB und Exzisionsbiopsie und IHC im Zeitraum			
	Zeitraum:	2002- 2006			
	Biopsieart	14g CNB, Anzahl Zylinder nicht angegeben			
	Test:	HercepTest (DAKO)			
Verblindeter Vergleich:	-				
Jacobs et al., 1998, USA	Erhebung	prospektiv	<u>Konkordanz ges:</u> <u>IHC:</u> 100%	IHC +: falls Membrananfärbung sichtbar war!	3b
	Studiengruppe:	n= 54 , konsekutiv, keine neoadj. Therapie			
	Zeitraum:	10/94-7/96			

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Biopsieart 23x CNB stereotaktisch, 14g, 27x CNB US gest., 13-18g, 6x CNB ohne Bildgebung; 2-6 Stanz-zylinder		IHC+ gesamt: 14,8% (8/54)	
	Test: IHC: CB11AK			
	Verblindeter Vergleich: nein			
Mann et al., 2004, Australien	Erhebung retrospektiv	<u>Konkordanz ges:</u>	CNB hat Tendenz zu höherem Scoring	3b
	Studiengruppe: n= 100 , konsekutiv, mit invasivem Tumor, mit CNB und FSE im Haus, ohne neoadj. Therapie	<u>IHC:</u> a) 80% (80/100) +/- b) 42% diff. 0-3+		
	Zeitraum: 4/99-3/2002			
	Biopsieart 14g CNB, US-gest.:	<u>Upgrade:</u> a) 18% b) 54% (meist 1+ CNB und 0 FSE)		
	Test: HercepTest (DAKO) beurteilt von 2 Pathologen und Arbitration bei <20% Differenz	<u>Downgrade:</u> a) 3% b) 4%		
	Verblindeter Vergleich: -	<u>Konkordant - (0/1+):</u> 77,8% (63/81) <u>Konkordant +/-(-2+):</u> 83,3% (10/12) <u>Konkordant 3+:</u> 100% (7/7)		
Müller- Holzner, 2001, Österreich	Erhebung -	<u>Konkordanz ges:</u>	CNB hat Tendenz zu höherem Scoring	3b
	Studiengruppe: n= 64 , T1-T4	<u>IHC:</u> a) 92,2% (59/64) +/- b) 81,3% (52/64) diff. nach 0-3+		
	Zeitraum: -			
	Biopsieart CNB 14g , US gest.	<u>Upgrade:</u> a) 5% b) 26,6% (meist 1+ CNB und 0 FSE)		
	Test: HercepTest (DAKO), mind. 3 Stanzzyylinder			

Autor/ Jahr/Land	Studienbeschreibung	Ergebnisse	Bemerkungen	Evidenz- grad
	Verblindeter Vergleich: -	Downgrade: a) 3% b) 6% <u>Konkordant 0:</u> 51,4% (18/35) <u>Konkordant 1+:</u> 83,3% (10/12) <u>Konkordant 2+:</u> 78,8% (7/9) <u>Konkordant 3+:</u> 89% (7/8)		
Taucher et al., 2004, Österreich	Erhebung retrospektiv Studiengruppe: n= 325 , mit ausreichend Biopsiematerial und FSE im Zeitraum, 110 Pat. mit neoadjuv. Chemotherapie, davon 13 mit CR- nicht weiter auswertbar, n= 312! Zeitraum: 1998-2002 Biopsieart 15g CNB 2-6 Stanzzyylinder Test: <i>HercepTest</i> (DAKO)+FISH (PathVysion) bei IHC 3+ Verblindeter Vergleich: -	<u>Konkordanz ges:</u> a) 92% (245/312) + /- K= 0.74 b) 96% (300/312) falls IHC- CNB 3+validiert durch FISH (K= 0.86) Konkordanz 0/1+/2+ (-): 96% 3+: 77,2% von 13 IHC 3+ CNB waren 11 FISH negativ Upgrade: a) 22,8% (13/57) (3+ CNB, - FSE) b) 4,3% (44/46) Downgrade: a) 4% (10/245) b) 3,9% (10/256)	IHC + nur bei 3+! Bei b) validiert durch FISH 110 Pat. mit neoadj. Therapie – hatte keinen Einfluss auf Her2-Status CNB Tendenz zu höherem Scoring	3b

5.5.2 Ausgeschlossene Studien im Vergleich Her-2/neu aus Stanzbiopsie - definitives Operationspräparat

Tabelle 10

Ausgeschlossene Studien im Vergleich Her-2/neu aus Stanzbiopsie - definitives Operationspräparat

Autor, Jahr, Land	Inhalt	Grund für Ausschluss
Gancberg et al., 2002, USA	Vergleich Her-2/neu in Tissue Microrarray und FSE	TMA aus FSE
Rody et al., 2006, BRD	Gen - Expressionprofile an Microarrays aus Stanzbiopsie vor neoadjuvanter Therapie – Korrelation mit Routine Prognosemarkern	TMA , nicht Original-Stanzbiopise
Kaneko et al., 2002, USA	<i>n= 46 Vergleich FISH aus touch imprints mit IHC aus FSE</i> <u>Konkordanz ges:</u> 91,3% (42/46) Downgrade FISH: 33,3% (2/6) Insgesamt 6x IHC+: 3x 3+; 3x 2+ 2x FISH negativ (1x bei 2+, 1x bei 3+)	Vergleich unterschiedlicher Verfahren, touch imprints kein Standard für FISH

6. Literatur

6.1 IHC und FISH - Ansprechen auf Trastuzumab

6.1.1 Eingeschlossene Studien zum Ansprechen auf Trastuzumab

1. Hofmann M, Stoss O, Gaiser T, Kneitz H, Heinmoller P, Gutjahr T, Kaufmann M, Henkel T, Ruschoff J. Central HER2 IHC and FISH analysis in a trastuzumab (Herceptin(R)) Phase II monotherapy study: assessment of test sensitivity and impact of chromosome 17 polysomy. *J Clin Pathol* 2007;
2. Mass RD, Press MF, Anderson S, Cobleigh MA, Vogel CL, Dybdal N, Leiberman G, Slamon DJ. Evaluation of clinical outcomes according to HER2 detection by fluorescence in situ hybridization in women with metastatic breast cancer treated with trastuzumab. *Clin Breast Cancer* 2005;6(3):240-6
3. Seidman AD, Fornier MN, Esteva FJ, Tan L, Kaptain S, Bach A, Panageas KS, Arroyo C, Valero V, Currie V, Gilewski T, Theodoulou M, Moynahan ME, Moasser M, Sklarin N, Dickler M, D'Andrea G, Cristofanilli M, Rivera E, Hortobagyi GN, Norton L, Hudis CA. Weekly trastuzumab and paclitaxel therapy for metastatic breast cancer with analysis of efficacy by HER2 immunophenotype and gene amplification. *J Clin Oncol* 2001;19(10):2587-95
4. Vogel CL, Cobleigh MA, Tripathy D, Gutheil JC, Harris LN, Fehrenbacher L, Slamon DJ, Murphy M, Novotny WF, Burchmore M, Shak S, Stewart SJ, Press M. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J Clin Oncol* 2002;20(3):719-26

6.1.2 Ausgeschlossene Studien zum Ansprechen auf Trastuzumab

1. Baselga J. Herceptin alone or in combination with chemotherapy in the treatment of HER2-positive metastatic breast cancer: pivotal trials. *Oncology* 2001;61 Suppl 2:14-21
2. Burstein HJ, Harris LN, Marcom PK, Lambert-Falls R, Havlin K, Overmoyer B, Friedlander RJ, Jr., Gargiulo J, Strenger R, Vogel CL, Ryan PD, Ellis MJ, Nunes RA, Bunnell CA, Campos SM, Hallor M, Gelman R, Winer EP. Trastuzumab and vinorelbine as first-line therapy for HER2-overexpressing metastatic breast cancer: multicenter phase II trial with clinical outcomes, analysis of serum tumor markers as predictive factors, and cardiac surveillance algorithm. *J Clin Oncol* 2003;21(15):2889-95
3. Diaz NM. Laboratory testing for HER2/neu in breast carcinoma: an evolving strategy to predict response to targeted therapy. *Cancer Control* 2001;8(5):415-8
4. Fornier MN, Seidman AD, Schwartz MK, Ghani F, Thiel R, Norton L, Hudis C. Serum HER2 extracellular domain in metastatic breast cancer patients treated with weekly trastuzumab and paclitaxel: association with HER2 status by immunohistochemistry and fluorescence in situ hybridization and with response rate. *Ann Oncol* 2005;16(2):234-9
5. Hudelist G, Kostler W, Czerwenka K, Kubista E, Singer CF. Predicting the clinical course of breast cancer patients undergoing trastuzumab-based therapy: an outlook. *Methods Find Exp Clin Pharmacol* 2004;26(3):201-10
6. Jahanzeb M. Trastuzumab-based combinations in metastatic breast cancer: how to make a choice. *Clin Breast Cancer* 2003;4(1):28-38
7. Leyland-Jones B. Maximizing the response to Herceptin therapy through optimal use and patient selection. *Anticancer Drugs* 2001;12 Suppl 4:S11-S17
8. Tedesco KL, Thor AD, Johnson DH, Shyr Y, Blum KA, Goldstein LJ, Gradishar WJ, Nicholson BP, Merkel DE, Murrey D, Edgerton S, Sledge GW, Jr. Docetaxel combined with trastuzumab is an active regimen in HER-2 3+ overexpressing and fluorescent in situ hybridization-positive metastatic breast cancer: a multi-institutional phase II trial. *J Clin Oncol* 2004;22(6):1071-7

9. Vogel CL, Cobleigh MA, Tripathy D, Gutheil JC, Harris LN, Fehrenbacher L, Slamon DJ, Murphy M, Novotny WF, Burchmore M, Shak S, Stewart SJ. First-line Herceptin monotherapy in metastatic breast cancer. *Oncology* 2001;61 Suppl 2:37-42

6.2. Reliabilität von IHC und FISH

6.2.1 Eingeschlossene Studien zur Reliabilität von IHC und FISH

1. Bartlett JM, Going JJ, Mallon EA, Watters AD, Reeves JR, Stanton P, Richmond J, Donald B, Ferrier R, Cooke TG. Evaluating HER2 amplification and overexpression in breast cancer. *J Pathol* 2001;195(4):422-8
2. Bilous M, Ades C, Armes J, Bishop J, Brown R, Cooke B, Cummings M, Farshid G, Field A, Morey A, McKenzie P, Raymond W, Robbins P, Tan L. Predicting the HER2 status of breast cancer from basic histopathology data: an analysis of 1500 breast cancers as part of the HER2000 International Study. *Breast* 2003;12(2):92-8
3. Dowsett M, Bartlett J, Ellis IO, Salter J, Hills M, Mallon E, Watters AD, Cooke T, Paish C, Wencyk PM, Pinder SE. Correlation between immunohistochemistry (HercepTest) and fluorescence in situ hybridization (FISH) for HER-2 in 426 breast carcinomas from 37 centres. *J Pathol* 2003;199(4):418-23
4. Dowsett M, Hanna WM, Kockx M, Penault-Llorca F, Ruschoff J, Gutjahr T, Habben K, van d, V. Standardization of HER2 testing: results of an international proficiency-testing ring study. *Mod Pathol* 2007;20(5):584-91
5. Dybdal N, Leiberman G, Anderson S, McCune B, Bajamonde A, Cohen RL, Mass RD, Sanders C, Press MF. Determination of HER2 gene amplification by fluorescence in situ hybridization and concordance with the clinical trials immunohistochemical assay in women with metastatic breast cancer evaluated for treatment with trastuzumab. *Breast Cancer Res Treat* 2005;93(1):3-11
6. Gancberg D, Jarvinen T, Di LA, Rouas G, Cardoso F, Paesmans M, Verhest A, Piccart MJ, Isola J, Larsimont D. Evaluation of HER-2/NEU protein expression in breast cancer by immunohistochemistry: an interlaboratory study assessing the reproducibility of HER-2/NEU testing. *Breast Cancer Res Treat* 2002;74(2):113-20
7. Hsu CY, Ho DM, Yang CF, Lai CR, Yu IT, Chiang H. Interobserver reproducibility of Her-2/neu protein overexpression in invasive breast carcinoma using the DAKO HercepTest. *Am J Clin Pathol* 2002;118(5):693-8
8. Italian Network for Quality Assurance of Tumor Biomarkers (INQAT) Group. Interobserver reproducibility of immunohistochemical HER-2/neu assessment in human breast cancer: an update from INQAT round III. *Int J Biol Markers* 2005;20(3):189-94
9. Lacroix-Triki M, Mathoulin-Pelissier S, Ghnassia JP, MacGrogan G, Vincent-Salomon A, Brouste V, Mathieu MC, Roger P, Bibeau F, Jacquemier J, Penault-Llorca F, Arnould L. High inter-observer agreement in immunohistochemical evaluation of HER-2/neu expression in breast cancer: a multicentre GEFPICS study. *Eur J Cancer* 2006;42(17):2946-53
10. Paik S, Bryant J, Tan-Chiu E, Romond E, Hiller W, Park K, Brown A, Yothers G, Anderson S, Smith R, Wickerham DL, Wolmark N. Real-world performance of HER2 testing--National Surgical Adjuvant Breast and Bowel Project experience. *J Natl Cancer Inst* 2002;94(11):852-4
11. Perez EA, Suman VJ, Davidson NE, Martino S, Kaufman PA, Lingle WL, Flynn PJ, Ingle JN, Visscher D, Jenkins RB. HER2 testing by local, central, and reference laboratories in specimens from the North Central Cancer Treatment Group N9831 intergroup adjuvant trial. *J Clin Oncol* 2006;24(19):3032-8
12. Press MF, Sauter G, Bernstein L, Villalobos IE, Mirlacher M, Zhou JY, Wardeh R, Li YT, Guzman R, Ma Y, Sullivan-Halley J, Santiago A, Park JM, Riva A, Slamon DJ. Diagnostic evaluation of HER-2 as a molecular target: an assessment of accuracy and reproducibility of laboratory testing in large, prospective, randomized clinical trials. *Clin Cancer Res* 2005;11(18):6598-607
13. Reddy JC, Reimann JD, Anderson SM, Klein PM. Concordance between central and local laboratory HER2 testing from a community-based clinical study. *Clin Breast Cancer* 2006;7(2):153-7

14. Rhodes A, Jasani B, Anderson E, Dodson AR, Balaton AJ. Evaluation of HER-2/neu immunohistochemical assay sensitivity and scoring on formalin-fixed and paraffin-processed cell lines and breast tumors: a comparative study involving results from laboratories in 21 countries. *Am J Clin Pathol* 2002;118(3):408-17
15. Roche PC, Suman VJ, Jenkins RB, Davidson NE, Martino S, Kaufman PA, Addo FK, Murphy B, Ingle JN, Perez EA. Concordance between local and central laboratory HER2 testing in the breast intergroup trial N9831. *J Natl Cancer Inst* 2002;94(11):855-7
16. Thomson TA, Hayes MM, Spinelli JJ, Hilland E, Sawrenko C, Phillips D, Dupuis B, Parker RL. HER-2/neu in breast cancer: interobserver variability and performance of immunohistochemistry with 4 antibodies compared with fluorescent in situ hybridization. *Mod Pathol* 2001;14(11):1079-86
17. Tsuda H, Akiyama F, Terasaki H, Hasegawa T, Kurosumi M, Shimadzu M, Yamamori S, Sakamoto G. Detection of HER-2/neu (c-erb B-2) DNA amplification in primary breast carcinoma. Interobserver reproducibility and correlation with immunohistochemical HER-2 overexpression. *Cancer* 2001;92(12):2965-74
18. Tsuda H, Sasano H, Akiyama F, Kurosumi M, Hasegawa T, Osamura RY, Sakamoto G. Evaluation of interobserver agreement in scoring immunohistochemical results of HER-2/neu (c-erbB-2) expression detected by HercepTest, Nichirei polyclonal antibody, CB11 and TAB250 in breast carcinoma. *Pathol Int* 2002;52(2):126-34
19. Yaziji H, Goldstein LC, Barry TS, Werling R, Hwang H, Ellis GK, Gralow JR, Livingston RB, Gown AM. HER-2 testing in breast cancer using parallel tissue-based methods. *JAMA* 2004;291(16):1972-7

6.2.2 Ausgeschlossene Studien zur Reliabilität von IHC und FISH

1. Italian Network for Quality Assurance of Tumor Biomarkers (INQAT) Group. Interobserver reproducibility of immunohistochemical HER-2/neu evaluation in human breast cancer: the real-world experience. *Int J Biol Markers* 2004;19(2):147-54
2. Gunhan O, Kafousi M, Kurt B, Koutsopoulos AV, Karslioglu Y, Stathopoulos E, Celasun B. Comparison of C-erbB2 (HER-2/Neu) immunohistochemistry results on invasive breast carcinomas. Experience of 2 Pathology departments in Turkey and Greece. *Anal Quant Cytol Histol* 2004;26(3):151-4

6.3. Konkordanz von IHC und FISH

6.3.1 Eingeschlossene Studien zur Konkordanz von IHC und FISH

1. Ainsworth R, Bartlett JM, Going JJ, Mallon EA, Forsyth A, Richmond J, Angerson W, Watters A, Dunne B. IHC for Her2 with CBE356 antibody is a more accurate predictor of Her2 gene amplification by FISH than HercepTest in breast carcinoma. *J Clin Pathol* 2005;58(10):1086-90
2. Bartlett JM, Going JJ, Mallon EA, Watters AD, Reeves JR, Stanton P, Richmond J, Donald B, Ferrier R, Cooke TG. Evaluating HER2 amplification and overexpression in breast cancer. *J Pathol* 2001;195(4):422-8
3. Bilous M, Ades C, Armes J, Bishop J, Brown R, Cooke B, Cummings M, Farshid G, Field A, Morey A, McKenzie P, Raymond W, Robbins P, Tan L. Predicting the HER2 status of breast cancer from basic histopathology data: an analysis of 1500 breast cancers as part of the HER2000 International Study. *Breast* 2003;12(2):92-8
4. Bouchalova K, Trojanec R, Kolar Z, Cwiertka K, Cernakova I, Mihal V, Hajduch M. Analysis of ERBB2 and TOP2A gene status using fluorescence in situ hybridization versus immunohistochemistry in localized breast cancer. *Neoplasma* 2006;53(5):393-401
5. Ciampa A, Xu B, Ayata G, Baiyee D, Wallace J, Wertheimer M, Edmiston K, Khan A. HER-2 status in breast cancer: correlation of gene amplification by FISH with immunohistochemistry expression using advanced cellular imaging system. *Appl Immunohistochem Mol Morphol* 2006;14(2):132-7

6. Cianciulli AM, Botti C, Coletta AM, Buglioni S, Marzano R, Benevolo M, Cione A, Mottolese M. Contribution of fluorescence in situ hybridization to immunohistochemistry for the evaluation of HER-2 in breast cancer. *Cancer Genet Cytogenet* 2002;133(1):66-71
7. Couturier J, Vincent-Salomon A, Nicolas A, Beuzeboc P, Mouret E, Zafrani B, Sastre-Garau X. Strong correlation between results of fluorescent in situ hybridization and immunohistochemistry for the assessment of the ERBB2 (HER-2/neu) gene status in breast carcinoma. *Mod Pathol* 2000;13(11):1238-43
8. Dendukuri N, Khetani K, Mclsaac M, Brophy J. Testing for HER2-positive breast cancer: a systematic review and cost-effectiveness analysis. *CMAJ* 2007;176(10):1429-34. Review
9. Dolan M, Snover D. Comparison of immunohistochemical and fluorescence in situ hybridization assessment of HER-2 status in routine practice. *Am J Clin Pathol* 2005;123(5):766-70
10. Dowsett M, Bartlett J, Ellis IO, Salter J, Hills M, Mallon E, Watters AD, Cooke T, Paish C, Wencyk PM, Pinder SE. Correlation between immunohistochemistry (HerceptTest) and fluorescence in situ hybridization (FISH) for HER-2 in 426 breast carcinomas from 37 centres. *J Pathol* 2003;199(4):418-23
11. Dybdal N, Leiberman G, Anderson S, McCune B, Bajamonde A, Cohen RL, Mass RD, Sanders C, Press MF. Determination of HER2 gene amplification by fluorescence in situ hybridization and concordance with the clinical trials immunohistochemical assay in women with metastatic breast cancer evaluated for treatment with trastuzumab. *Breast Cancer Res Treat* 2005;93(1):3-11
12. Ellis CM, Dyson MJ, Stephenson TJ, Maltby EL. HER2 amplification status in breast cancer: a comparison between immunohistochemical staining and fluorescence in situ hybridisation using manual and automated quantitative image analysis scoring techniques. *J Clin Pathol* 2005;58(7):710-4
13. Gancberg D, Lespagnard L, Rouas G, Paesmans M, Piccart M, Di LA, Nogaret JM, Hertens D, Verhest A, Larsimont D. Sensitivity of HER-2/neu antibodies in archival tissue samples of invasive breast carcinomas. Correlation with oncogene amplification in 160 cases. *Am J Clin Pathol* 2000;113(5):675-82
14. Garcia-Caballero T, Menendez MD, Vazquez-Boquete A, Gallego R, Forteza J, Fraga M. HER-2 status determination in breast carcinomas. A practical approach. *Histol Histopathol* 2006;21(3):227-36. Ref ID: 8987
15. Hanna WM, Kwok K. Chromogenic in-situ hybridization: a viable alternative to fluorescence in-situ hybridization in the HER2 testing algorithm. *Mod Pathol* 2006;19(4):481-7
16. Harris LN, Liotcheva V, Broadwater G, Ramirez MJ, Maimonis P, Anderson S, Everett T, Harpole D, Moore MB, Berry DA, Rizzeri D, Vredenburgh JJ, Bentley RC. Comparison of methods of measuring HER-2 in metastatic breast cancer patients treated with high-dose chemotherapy. *J Clin Oncol* 2001;19(6):1698-706
17. Hoang MP, Sahin AA, Ordonez NG, Sneige N. HER-2/neu gene amplification compared with HER-2/neu protein overexpression and interobserver reproducibility in invasive breast carcinoma. *Am J Clin Pathol* 2000;113(6):852-9
18. Jacobs TW, Gown AM, Yaziji H, Barnes MJ, Schnitt SJ. Comparison of fluorescence in situ hybridization and immunohistochemistry for the evaluation of HER-2/neu in breast cancer. *J Clin Oncol* 1999;17(7):1974-82
19. Kakar S, Puangsuwan N, Stevens JM, Serenas R, Mangan G, Sahai S, Mihalov ML. HER-2/neu assessment in breast cancer by immunohistochemistry and fluorescence in situ hybridization: comparison of results and correlation with survival. *Mol Diagn* 2000;5(3):199-207
20. Lal P, Salazar PA, Hudis CA, Ladanyi M, Chen B. HER-2 testing in breast cancer using immunohistochemical analysis and fluorescence in situ hybridization: a single-institution experience of 2,279 cases and comparison of dual-color and single-color scoring. *Am J Clin Pathol* 2004;121(5):631-6
21. Lan C, Liu JM, Liu TW, Hsu DH, Liang S, Chen JR, Peng JW. erb-b2 amplification by fluorescence in situ hybridization in breast cancer specimens read as 2+ in immunohistochemical analysis. *Am J Clin Pathol* 2005;124(1):97-102
22. Larsimont D, Di LA, Rouas G, Paesmans M, Ferreira-Filho F, Bernard C, Cardoso F, Verhest A, Piccart MJ, Gancberg D. HER-2/neu evaluation by immunohistochemistry and fluorescence in situ hybridization in breast cancer: implications for daily laboratory practice. *Anticancer Res* 2002;22(4):2485-90

23. Lebeau A, Deimling D, Kaltz C, Sendelhofert A, Iff A, Luthardt B, Untch M, Lohrs U. Her-2/neu analysis in archival tissue samples of human breast cancer: comparison of immunohistochemistry and fluorescence in situ hybridization. *J Clin Oncol* 2001;19(2):354-63
24. Lopez-Guerrero JA, Navarro S, Noguera R, Almenar S, Pellin A, Vazquez C, Llombart-Bosch A. Histological tumor grade correlates with HER2/c-erbB-2 status in invasive breast cancer: a comparative analysis between immunohistochemical (CB11 clone and Herceptest), FISH and differential PCR procedures. *Arkh Patol* 2003;65(1):50-5
25. Loring P, Cummins R, O'Grady A, Kay EW. HER2 positivity in breast carcinoma: a comparison of chromogenic in situ hybridization with fluorescence in situ hybridization in tissue microarrays, with targeted evaluation of intratumoral heterogeneity by in situ hybridization. *Appl Immunohistochem Mol Morphol* 2005;13(2):194-200
26. Lottner C, Schwarz S, Diermeier S, Hartmann A, Knuechel R, Hofstaedter F, Brockhoff G. Simultaneous detection of HER2/neu gene amplification and protein overexpression in paraffin-embedded breast cancer. *J Pathol* 2005;205(5):577-84
27. McCormick SR, Lillemoie TJ, Beneke J, Schrauth J, Reinartz J. HER2 assessment by immunohistochemical analysis and fluorescence in situ hybridization: comparison of HercepTest and PathVysion commercial assays. *Am J Clin Pathol* 2002;117(6):935-43
28. Mrozkowiak A, Olszewski WP, Piascik A, Olszewski WT. HER2 status in breast cancer determined by IHC and FISH: comparison of the results. *Pol J Pathol* 2004;55(4):165-71
29. Ogura H, Akiyama F, Kasumi F, Kazui T, Sakamoto G. Evaluation of HER-2 status in breast carcinoma by fluorescence in situ hybridization and immunohistochemistry. *Breast Cancer* 2003;10(3):234-40
30. Owens MA, Horten BC, Da Silva MM. HER2 amplification ratios by fluorescence in situ hybridization and correlation with immunohistochemistry in a cohort of 6556 breast cancer tissues. *Clin Breast Cancer* 2004;5(1):63-9
31. Paik S, Bryant J, Tan-Chiu E, Romond E, Hiller W, Park K, Brown A, Yothers G, Anderson S, Smith R, Wickerham DL, Wolmark N. Real-world performance of HER2 testing--National Surgical Adjuvant Breast and Bowel Project experience. *J Natl Cancer Inst* 2002;94(11):852-4
32. Pauletti G, Dandekar S, Rong H, Ramos L, Peng H, Seshadri R, Slamon DJ. Assessment of methods for tissue-based detection of the HER-2/neu alteration in human breast cancer: a direct comparison of fluorescence in situ hybridization and immunohistochemistry. *J Clin Oncol* 2000;18(21):3651-64
33. Perez EA, Roche PC, Jenkins RB, Reynolds CA, Halling KC, Ingle JN, Wold LE. HER2 testing in patients with breast cancer: poor correlation between weak positivity by immunohistochemistry and gene amplification by fluorescence in situ hybridization. *Mayo Clin Proc* 2002;77(2):148-54
34. Press MF, Slamon DJ, Flom KJ, Park J, Zhou JY, Bernstein L. Evaluation of HER-2/neu gene amplification and overexpression: comparison of frequently used assay methods in a molecularly characterized cohort of breast cancer specimens. *J Clin Oncol* 2002;20(14):3095-105
35. Press MF, Sauter G, Bernstein L, Villalobos IE, Mirlacher M, Zhou JY, Wardeh R, Li YT, Guzman R, Ma Y, Sullivan-Halley J, Santiago A, Park JM, Riva A, Slamon DJ. Diagnostic evaluation of HER-2 as a molecular target: an assessment of accuracy and reproducibility of laboratory testing in large, prospective, randomized clinical trials. *Clin Cancer Res* 2005;11(18):6598-607
36. Ridolfi RL, Jamehdor MR, Arber JM. HER-2/neu testing in breast carcinoma: a combined immunohistochemical and fluorescence in situ hybridization approach. *Mod Pathol* 2000;13(8):866-73
37. Roche PC, Suman VJ, Jenkins RB, Davidson NE, Martino S, Kaufman PA, Addo FK, Murphy B, Ingle JN, Perez EA. Concordance between local and central laboratory HER2 testing in the breast intergroup trial N9831. *J Natl Cancer Inst* 2002;94(11):855-7
38. Rossi E, Ubiali A, Cadei M, Balzarini P, Valagussa E, Lucini L, Alpi F, Galletti A, Fontana L, Tedoldi C, Grigolato P. HER-2/neu in breast cancer: a comparative study between histology, immunohistochemistry, and molecular technique (FISH). *Appl Immunohistochem Mol Morphol* 2006;14(2):127-31

39. Saez A, Andreu FJ, Segui MA, Bare ML, Fernandez S, Dinares C, Rey M. HER-2 gene amplification by chromogenic in situ hybridisation (CISH) compared with fluorescence in situ hybridisation (FISH) in breast cancer-A study of two hundred cases. *Breast* 2006;15(4):519-27
40. Sauer T, Wiedswang G, Boudjema G, Christensen H, Karesen R. Assessment of HER-2/neu overexpression and/or gene amplification in breast carcinomas: should in situ hybridization be the method of choice? *APMIS* 2003;111(3):444-50
41. Seidman AD, Fournier MN, Esteva FJ, Tan L, Kaptain S, Bach A, Panageas KS, Arroyo C, Valero V, Currie V, Gilewski T, Theodoulou M, Moynahan ME, Moasser M, Sklarin N, Dickler M, D'Andrea G, Cristofanilli M, Rivera E, Hortobagyi GN, Norton L, Hudis CA. Weekly trastuzumab and paclitaxel therapy for metastatic breast cancer with analysis of efficacy by HER2 immunophenotype and gene amplification. *J Clin Oncol* 2001;19(10):2587-95
42. Tanner M, Gancberg D, Di LA, Larsimont D, Rouas G, Piccart MJ, Isola J. Chromogenic in situ hybridization: a practical alternative for fluorescence in situ hybridization to detect HER-2/neu oncogene amplification in archival breast cancer samples. *Am J Pathol* 2000;157(5):1467-72
43. Tawfik OW, Kimler BF, Davis M, Donahue JK, Persons DL, Fan F, Hagemeister S, Thomas P, Connor C, Jewell W, Fabian CJ. Comparison of immunohistochemistry by automated cellular imaging system (ACIS) versus fluorescence in-situ hybridization in the evaluation of HER-2/neu expression in primary breast carcinoma. *Histopathology* 2006;48(3):258-67
44. Tubbs RR, Pettay JD, Roche PC, Stoler MH, Jenkins RB, Grogan TM. Discrepancies in clinical laboratory testing of eligibility for trastuzumab therapy: apparent immunohistochemical false-positives do not get the message. *J Clin Oncol* 2001;19(10):2714-21
45. Vocaturo A, Novelli F, Benevolo M, Piperno G, Marandino F, Cianciulli AM, Merola R, Donnorso RP, Sperduti I, Buglioni S, Mottolese M. Chromogenic in situ hybridization to detect HER-2/neu gene amplification in histological and ThinPrep-processed breast cancer fine-needle aspirates: a sensitive and practical method in the trastuzumab era. *Oncologist* 2006;11(8):878-86
46. Wang S, Saboorian MH, Frenkel EP, Haley BB, Siddiqui MT, Gokaslan S, Wians FH, Jr., Hynan L, Ashfaq R. Assessment of HER-2/neu status in breast cancer. Automated Cellular Imaging System (ACIS)-assisted quantitation of immunohistochemical assay achieves high accuracy in comparison with fluorescence in situ hybridization assay as the standard. *Am J Clin Pathol* 2001;116(4):495-503
47. Yaziji H, Goldstein LC, Barry TS, Werling R, Hwang H, Ellis GK, Galow JR, Livingston RB, Gown AM. HER-2 testing in breast cancer using parallel tissue-based methods. *JAMA* 2004;291(16):1972-7
48. Zhao J, Wu R, Au A, Marquez A, Yu Y, Shi Z. Determination of HER2 gene amplification by chromogenic in situ hybridization (CISH) in archival breast carcinoma. *Mod Pathol* 2002;15(6):657-65.

6.3.2. Ausgeschlossene Studien zur Konkordanz von IHC und FISH

1. Arnould L, Denoux Y, MacGrogan G, Penault-Llorca F, Fiche M, Treilleux I, Mathieu MC, Vincent-Salomon A, Vilain MO, Couturier J. Agreement between chromogenic in situ hybridisation (CISH) and FISH in the determination of HER2 status in breast cancer. *Br J Cancer* 2003;88(10):1587-91
2. Bankfalvi A, Simon R, Brandt B, Burger H, Vollmer I, Dockhorn-Dworniczak B, Lelle RJ, Boecker W. Comparative methodological analysis of erbB-2/HER-2 gene dosage, chromosomal copy number and protein overexpression in breast carcinoma tissues for diagnostic use. *Histopathology* 2000;37(5):411-9
3. Bilous M, Morey A, Armes J, Cummings M, Francis G. Chromogenic in situ hybridisation testing for HER2 gene amplification in breast cancer produces highly reproducible results concordant with fluorescence in situ hybridisation and immunohistochemistry. *Pathology* 2006;38(2):120-4
4. Bussolati G, Montemurro F, Righi L, Donadio M, Aglietta M, Sapino A. A modified Trastuzumab antibody for the immunohistochemical detection of HER-2 overexpression in breast cancer. *Br J Cancer* 2005;92(7):1261-7
5. Choi DH, Shin DB, Lee MH, Lee DW, Dhandapani D, Carter D, King BL, Haffty BG. A comparison of five immunohistochemical biomarkers and HER-2/neu gene amplification by fluorescence in situ

- hybridization in white and Korean patients with early-onset breast carcinoma. *Cancer* 2003;98(8):1587-95
6. Dagrada GP, Mezzelani A, Alasio L, Ruggeri M, Romano R, Pierotti MA, Pilotti S. HER-2/neu assessment in primary chemotherapy treated breast carcinoma: no evidence of gene profile changing. *Breast Cancer Res Treat* 2003;80(2):207-14
 7. Dandachi N, Dietze O, Hauser-Kronberger C. Evaluation of the clinical significance of HER2 amplification by chromogenic in situ hybridisation in patients with primary breast cancer. *Anticancer Res* 2004;24(4):2401-6
 8. Diaz LK, Gupta R, Kidwai N, Sneige N, Wiley EL. The use of TMA for interlaboratory validation of FISH testing for detection of HER2 gene amplification in breast cancer. *J Histochem Cytochem* 2004;52(4):501-7
 9. Di Palma S, Collins N, Faulkes C, Ping B, Ferns G, Haagsma B, Kissin M, Layer G, Cook M. Chromogenic in-situ hybridisation (CISH) should be an accepted method in the routine diagnostic evaluation of HER2 status in breast cancer. *J Clin Pathol* 2007
 10. Gancberg D, Lespagnard L, Rouas G, Paesmans M, Piccart M, Di LA, Nogaret JM, Hertens D, Verhest A, Larsimont D. Sensitivity of HER-2/neu antibodies in archival tissue samples of invasive breast carcinomas. Correlation with oncogene amplification in 160 cases. *Am J Clin Pathol* 2000;113(5):675-82
 11. Ginestier C, Charafe-Jauffret E, Penault-Llorca F, Geneix J, Adelaide J, Chaffanet M, Mozziconacci MJ, Hassoun J, Viens P, Birnbaum D, Jacquemier J. Comparative multi-methodological measurement of ERBB2 status in breast cancer. *J Pathol* 2004;202(3):286-98
 12. Gouvea AP, Milanezi F, Olson SJ, Leitao D, Schmitt FC, Gobbi H. Selecting antibodies to detect HER2 overexpression by immunohistochemistry in invasive mammary carcinomas. *Appl Immunohistochem Mol Morphol* 2006;14(1):103-8
 13. Harbeck N, Ross JS, Yurdseven S, Dettmar P, Polcher M, Kuhn W, Ulm K, Graeff H, Schmitt M. HER-2/neu gene amplification by fluorescence in situ hybridization allows risk-group assessment in node-negative breast cancer. *Int J Oncol* 1999;14(4):663-71
 14. Hatanaka Y, Hashizume K, Kamihara Y, Itoh H, Tsuda H, Osamura RY, Tani Y. Quantitative immunohistochemical evaluation of HER2/neu expression with HercepTest™ in breast carcinoma by image analysis. *Pathol Int* 2001;51(1):33-6
 15. Jacobs TW, Gown AM, Yaziji H, Barnes MJ, Schnitt SJ. Related Articles, Links Specificity of HercepTest in determining HER-2/neu status of breast cancers using the United States Food and Drug Administration-approved scoring system. *J Clin Oncol* 1999;(7):1983-7
 16. Jimenez RE, Wallis T, Tabasczka P, Visscher DW. Determination of Her-2/Neu status in breast carcinoma: comparative analysis of immunohistochemistry and fluorescent in situ hybridization. *Mod Pathol* 2000;13(1):37-45
 17. Kakar S, Puangsuwan N, Stevens JM, Serenas R, Mangan G, Sahai S, Mihalov ML. Comparison of PathVysion and INFORM fluorescence in situ hybridization kits for assessment of HER-2/neu status in breast carcinoma. *Mol Diagn* 2000;5(3):193-7
 18. Kobayashi M, Ooi A, Oda Y, Nakanishi I. Protein overexpression and gene amplification of c-erbB-2 in breast carcinomas: a comparative study of immunohistochemistry and fluorescence in situ hybridization of formalin-fixed, paraffin-embedded tissues. *Hum Pathol* 2002;33(1):21-8
 19. Kunitomo K, Takehana T, Inoue S, Fujii H, Suzuki S, Matsumoto Y, Ooi A. Detection of c-erbB-2 (HER-2/neu) amplification in breast carcinoma by fluorescence in situ hybridization on tissue sections and imprinted cells. *Pathol Int* 2002;52(7):451-7
 20. Li Ning TE, Ronchetti R, Torres-Cabala C, Merino MJ. Role of chromogenic in situ hybridization (CISH) in the evaluation of HER2 status in breast carcinoma: comparison with immunohistochemistry and FISH. *Int J Surg Pathol* 2005;13(4):343-51
 21. Merkelbach-Bruse S, Wardelmann E, Behrens P, Losen I, Buettner R, Friedrichs N. Current diagnostic methods of HER-2/neu detection in breast cancer with special regard to real-time PCR. *Am J Surg Pathol* 2003;27(12):1565-70

22. O'Malley FP, Parkes R, Latta E, Tjan S, Zadro T, Mueller R, Arneson N, Blackstein M, Andrulis I. Comparison of HER2/neu status assessed by quantitative polymerase chain reaction and immunohistochemistry. *Am J Clin Pathol* 2001;115(4):504-11
23. Peiro G, Aranda FI, Adrover E, Niveiro M, Alenda C, Paya A, Segui J. Analysis of HER2 by chromogenic in situ hybridization and immunohistochemistry in lymph node-negative breast carcinoma: Prognostic relevance. *Hum Pathol* 2007;38(1):26-34
24. Perez EA, Suman VJ, Davidson NE, Martino S, Kaufman PA, Lingle WL, Flynn PJ, Ingle JN, Visscher D, Jenkins RB. HER2 testing by local, central, and reference laboratories in specimens from the North Central Cancer Treatment Group N9831 intergroup adjuvant trial. *J Clin Oncol* 2006;24(19):3032-8
25. Sapino A, Coccorullo Z, Cassoni P, Ghisolfi G, Gugliotta P, Bongiovanni M, Arisio R, Crafa P, Bussolati G. Which breast carcinomas need HER-2/neu gene study after immunohistochemical analysis? Results of combined use of antibodies against different c-erbB2 protein domains. *Histopathology* 2003;43(4):354-62
26. Selvarajan S, Bay BH, Chng MJ, Tan PH. The HercepTest and routine C-erbB2 immunohistochemistry in breast cancer: any difference? *Ann Acad Med Singapore* 2004;33(4):473-6
27. Sidoni A, Ferri I, Cavaliere A, Bellezza G, Scheibel M, Bucciarelli E. Detection of HER-2/neu (c-erbB-2) overexpression and amplification in breast carcinomas with ambiguous immunohistochemical results. A further contribution to defining the role of fluorescent in situ hybridization. *Anticancer Res* 2006;26(3B):2333-7
28. Tapia C, Schraml P, Simon R, Al-Kuraya KS, Maurer R, Mirlacher M, Novotny H, Spichtin H, Mihatsch MJ, Sauter G. HER2 analysis in breast cancer: reduced immunoreactivity in FISH non-informative cancer biopsies. *Int J Oncol* 2004;25(6):1551-7
29. Thomson TA, Hayes MM, Spinelli JJ, Hilland E, Sawrenko C, Phillips D, Dupuis B, Parker RL. HER-2/neu in breast cancer: interobserver variability and performance of immunohistochemistry with 4 antibodies compared with fluorescent in situ hybridization. *Mod Pathol* 2001;14(11):1079-86
30. Tsuda H, Tani Y, Hasegawa T, Fukutomi T. Concordance in judgments among c-erbB-2 (HER2/neu) overexpression detected by two immunohistochemical tests and gene amplification detected by Southern blot hybridization in breast carcinoma. *Pathol Int* 2001;51(1):26-32
31. Tubbs R, Pettay J, Skacel M, Powell R, Stoler M, Roche P, Hainfeld J. Gold-facilitated in situ hybridization: a bright-field autometallographic alternative to fluorescence in situ hybridization for detection of Her-2/neu gene amplification. *Am J Pathol* 2002;160(5):1589-95
32. Wixom CR, Albers EA, Weidner N. Her2 amplification: correlation of chromogenic in situ hybridization with immunohistochemistry and fluorescence in situ hybridization. *Appl Immunohistochem Mol Morphol* 2004;12(3):248-51

6.4. Konkordanz und Reliabilität FISH-CISH

6.4.1 Eingeschlossene Studien zu FISH-CISH

1. Arnould L, Denoux Y, MacGrogan G, Penault-Llorca F, Fiche M, Treilleux I, Mathieu MC, Vincent-Salomon A, Vilain MO, Couturier J. Agreement between chromogenic in situ hybridisation (CISH) and FISH in the determination of HER2 status in breast cancer. *Br J Cancer* 2003;88(10):1587-91
2. Bhargava R, Lal P, Chen B. Chromogenic in situ hybridization for the detection of HER-2/neu gene amplification in breast cancer with an emphasis on tumors with borderline and low-level amplification: does it measure up to fluorescence in situ hybridization? *Am J Clin Pathol* 2005;123(2):237-43
3. Bilous M, Morey A, Armes J, Cummings M, Francis G. Chromogenic in situ hybridisation testing for HER2 gene amplification in breast cancer produces highly reproducible results concordant with fluorescence in situ hybridisation and immunohistochemistry. *Pathology* 2006;38(2):120-4
4. Dandachi N, Dietze O, Hauser-Kronberger C. Evaluation of the clinical significance of HER2 amplification by chromogenic in situ hybridisation in patients with primary breast cancer. *Anticancer Res* 2004;24(4):2401-6

5. Di Palma S, Collins N, Faulkes C, Ping B, Ferns G, Haagsma B, Kissin M, Layer G, Cook M. Chromogenic in-situ hybridisation (CISH) should be an accepted method in the routine diagnostic evaluation of HER2 status in breast cancer. *J Clin Pathol* 2007;
6. Diaz LK, Gupta R, Kidwai N, Sneige N, Wiley EL. The use of TMA for interlaboratory validation of FISH testing for detection of HER2 gene amplification in breast cancer. *J Histochem Cytochem* 2004;52(4):501-7
7. Gong Y, Gilcrease M, Sneige N. Reliability of chromogenic in situ hybridization for detecting HER-2 gene status in breast cancer: comparison with fluorescence in situ hybridization and assessment of interobserver reproducibility. *Mod Pathol* 2005;18(8):1015-21
8. Gupta D, Middleton LP, Whitaker MJ, Abrams J. Comparison of fluorescence and chromogenic in situ hybridization for detection of HER-2/neu oncogene in breast cancer. *Am J Clin Pathol* 2003;119(3):381-7
9. Hanna WM, Kwok K. Chromogenic in-situ hybridization: a viable alternative to fluorescence in-situ hybridization in the HER2 testing algorithm. *Mod Pathol* 2006;19(4):481-7
10. Isola J, Tanner M, Forsyth A, Cooke TG, Watters AD, Bartlett JM. Interlaboratory comparison of HER-2 oncogene amplification as detected by chromogenic and fluorescence in situ hybridization. *Clin Cancer Res* 2004;10(14):4793-8
11. Laakso M, Tanner M, Isola J. Dual-colour chromogenic in situ hybridization for testing of HER-2 oncogene amplification in archival breast tumours. *J Pathol* 2006;210(1):3-9
12. Li Ning TE, Ronchetti R, Torres-Cabala C, Merino MJ. Role of chromogenic in situ hybridization (CISH) in the evaluation of HER2 status in breast carcinoma: comparison with immunohistochemistry and FISH. *Int J Surg Pathol* 2005;13(4):343-51
13. Loring P, Cummins R, O'Grady A, Kay EW. HER2 positivity in breast carcinoma: a comparison of chromogenic in situ hybridization with fluorescence in situ hybridization in tissue microarrays, with targeted evaluation of intratumoral heterogeneity by in situ hybridization. *Appl Immunohistochem Mol Morphol* 2005;13(2):194-200
14. Park K, Kim J, Lim S, Han S, Lee JY. Comparing fluorescence in situ hybridization and chromogenic in situ hybridization methods to determine the HER2/neu status in primary breast carcinoma using tissue microarray. *Mod Pathol* 2003;16(9):937-43
15. Peiro G, Aranda FI, Adrover E, Niveiro M, Alenda C, Paya A, Segui J. Analysis of HER2 by chromogenic in situ hybridization and immunohistochemistry in lymph node-negative breast carcinoma: Prognostic relevance. *Hum Pathol* 2007;38(1):26-34
16. Saez A, Andreu FJ, Segui MA, Bare ML, Fernandez S, Dinares C, Rey M. HER-2 gene amplification by chromogenic in situ hybridisation (CISH) compared with fluorescence in situ hybridisation (FISH) in breast cancer-A study of two hundred cases. *Breast* 2006;15(4):519-27
17. Tanner M, Gancberg D, Di LA, Larsimont D, Rouas G, Piccart MJ, Isola J. Chromogenic in situ hybridization: a practical alternative for fluorescence in situ hybridization to detect HER-2/neu oncogene amplification in archival breast cancer samples. *Am J Pathol* 2000;157(5):1467-72
18. Vocaturo A, Novelli F, Benevolo M, Piperno G, Marandino F, Cianciulli AM, Merola R, Donnorso RP, Sperduti I, Buglioni S, Mottolese M. Chromogenic in situ hybridization to detect HER-2/neu gene amplification in histological and ThinPrep-processed breast cancer fine-needle aspirates: a sensitive and practical method in the trastuzumab era. *Oncologist* 2006;11(8):878-86
19. Wixom CR, Albers EA, Weidner N. Her2 amplification: correlation of chromogenic in situ hybridization with immunohistochemistry and fluorescence in situ hybridization. *Appl Immunohistochem Mol Morphol* 2004;12(3):248-51
20. Zhao J, Wu R, Au A, Marquez A, Yu Y, Shi Z. Determination of HER2 gene amplification by chromogenic in situ hybridization (CISH) in archival breast carcinoma. *Mod Pathol* 2002;15(6):657-65

6.5 Übereinstimmung Biopsie – definitives Operationspräparat

6.5.1 Eingeschlossene Studien zur Übereinstimmung Biopsie – definitives Operationspräparat

1. Burge CN, Chang HR, Apple SK. Do the histologic features and results of breast cancer biomarker studies differ between core biopsy and surgical excision specimens? *Breast* 2006;15(2):167-72
2. Cahill RA, Walsh D, Landers RJ, Watson RG. Preoperative profiling of symptomatic breast cancer by diagnostic core biopsy. *Ann Surg Oncol* 2006;13(1):45-51
3. Cavaliere A, Sidoni A, Scheibel M, Bellezza G, Brachelente G, Vitali R, Bucciarelli E. Biopathologic profile of breast cancer core biopsy: is it always a valid method? *Cancer Lett* 2005;218(1):117-21
4. Connor CS, Tawfik OW, Joyce AJ, Davis MK, Mayo MS, Jewell WR. A comparison of prognostic tumor markers obtained on image-guided breast biopsies and final surgical specimens. *Am J Surg* 2002;184(4):322-4
5. Erggelet J, Große R, Holzhausen HJ, Hauptmann S, Thomssen C. Correlation of human epidermal growth factor receptor 2 (HER2), estrogen receptor (ER), and progesterone receptor (PR) expression as predicted by core biopsy with the immunohistochemical results of surgical breast cancer specimens. *Breast Care* 2007;2(2):94-8
6. Jacobs TW, Siziopikou KP, Prioleau JE, Raza S, Baum JK, Hayes DF, Schnitt SJ. Do prognostic marker studies on core needle biopsy specimens of breast carcinoma accurately reflect the marker status of the tumor? *Mod Pathol* 1998;11(3):259-64
7. Mann GB, Fahey VD, Feleppa F, Buchanan MR. Reliance on hormone receptor assays of surgical specimens may compromise outcome in patients with breast cancer. *J Clin Oncol* 2005;23(22):5148-54
8. Mueller-Holzner E, Fink V, Frede T, Marth C. Immunohistochemical determination of HER2 expression in breast cancer from core biopsy specimens: a reliable predictor of HER2 status of the whole tumor. *Breast Cancer Res Treat* 2001;69(1):13-9
9. Taucher S, Rudas M, Mader RM, Gnant M, Dubsy P, Roka S, Bachleitner T, Kandioler D, Steger G, Mittlbock M, Jakesz R. Prognostic markers in breast cancer: the reliability of HER2/neu status in core needle biopsy of 325 patients with primary breast cancer. *Wien Klin Wochenschr* 2004;116(1-2):26-31

6.5.2 Ausgeschlossene Studien zur Übereinstimmung Biopsie – definitives Operationspräparat

1. Gancberg D, Di Leo A, Rouas G, Jarvinen T, Verhest A, Isola J, Piccart MJ, Larsimont D. Reliability of the tissue microarray based FISH for evaluation of the HER-2 oncogene in breast carcinoma. *J Clin Pathol* 2002;55(4):315-7
2. Kaneko S, Gerasimova T, Butler WM, Cupples TE, Guerry PL, Greene GR, Young SR. The use of FISH on breast core needle samples for the presurgical assessment of HER-2 oncogene status. *Exp Mol Pathol* 2002;73(1):61-6
3. Rody A, Karn T, Gatje R, Kourtis K, Minckwitz G, Loibl S, Munnes M, Ruckhaberle E, Holtrich U, Kaufmann M, Ahr A. Gene expression profiles of breast cancer obtained from core cut biopsies before neoadjuvant docetaxel, adriamycin, and cyclophosphamide chemotherapy correlate with routine prognostic markers and could be used to identify predictive signatures. *Zentralbl Gynakol*

Anhang 1: Oxford Levels of Evidence

Oxford Centre for Evidence-based Medicine Levels of Evidence (May 2001)

Level	Therapy/Prevention, Aetiology/Harm	Prognosis	Diagnosis	Differential diagnosis/symptom prevalence study	Economic and decision analyses
1a	SR (with homogeneity*) of RCTs	SR (with homogeneity*) of inception cohort studies; CDR† validated in different populations	SR (with homogeneity*) of Level 1 diagnostic studies; CDR† with 1b studies from different clinical centres	SR (with homogeneity*) of prospective cohort studies	SR (with homogeneity*) of Level 1 economic studies
1b	Individual RCT (with narrow Confidence Interval‡)	Individual inception cohort study with ≥ 80% follow-up; CDR† validated in a single population	Validating** cohort study with good††† reference standards; or CDR† tested within one clinical centre	Prospective cohort study with good follow-up****	Analysis based on clinically sensible costs or alternatives; systematic review(s) of the evidence; and including multi-way sensitivity analyses
1c	All or none§	All or none case-series	Absolute SpPins and SnNouts††	All or none case-series	Absolute better-value or worse-value analyses ††††
2a	SR (with homogeneity*) of cohort studies	SR (with homogeneity*) of either retrospective cohort studies or untreated control groups in RCTs	SR (with homogeneity*) of Level >2 diagnostic studies	SR (with homogeneity*) of 2b and better studies	SR (with homogeneity*) of Level >2 economic studies
2b	Individual cohort study (including low quality RCT; e.g., <80% follow-up)	Retrospective cohort study or follow-up of untreated control patients in an RCT; Derivation of CDR† or validated on split-sample§§§ only	Exploratory** cohort study with good†††reference standards; CDR† after derivation, or validated only on split-sample§§§ or databases	Retrospective cohort study, or poor follow-up	Analysis based on clinically sensible costs or alternatives; limited review(s) of the evidence, or single studies; and including multi-way sensitivity analyses
2c	"Outcomes" Research; Ecological studies	"Outcomes" Research		Ecological studies	Audit or outcomes research
3a	SR (with homogeneity*) of case-control studies		SR (with homogeneity*) of 3b and better studies	SR (with homogeneity*) of 3b and better studies	SR (with homogeneity*) of 3b and better studies
3b	Individual Case-Control Study		Non-consecutive study; or without consistently applied reference standards	Non-consecutive cohort study, or very limited population	Analysis based on limited alternatives or costs, poor quality estimates of data, but including sensitivity analyses incorporating clinically sensible variations.
4	Case-series (and poor quality cohort and case-control studies§§)	Case-series (and poor quality prognostic cohort studies****)	Case-control study, poor or non-independent reference standard	Case-series or superseded reference standards	Analysis with no sensitivity analysis
5	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on economic theory or "first principles"

Produced by Bob Phillips, Chris Ball, Dave Sackett, Doug Badenoch, Sharon Straus, Brian Haynes, Martin Dawes since November 1998.

Notes

Users can add a minus-sign "-" to denote the level of that fails to provide a conclusive answer because of:

- EITHER a single result with a wide Confidence Interval (such that, for example, an ARR in an RCT is not statistically significant but whose confidence intervals fail to exclude clinically important benefit or harm)
- OR a Systematic Review with troublesome (and statistically significant) heterogeneity.
- Such evidence is inconclusive, and therefore can only generate Grade D recommendations.

*	By homogeneity we mean a systematic review that is free of worrisome variations (heterogeneity) in the directions and degrees of results between individual studies. Not all systematic reviews with statistically significant heterogeneity need be worrisome, and not all worrisome heterogeneity need be statistically significant. As noted above, studies displaying worrisome heterogeneity should be tagged with a "-" at the end of their designated level.
†	Clinical Decision Rule. (These are algorithms or scoring systems which lead to a prognostic estimation or a diagnostic category.)
‡	See note #2 for advice on how to understand, rate and use trials or other studies with wide confidence intervals.
§	Met when <u>all</u> patients died before the Rx became available, but some now survive on it; or when some patients died before the Rx became available, but <u>none</u> now die on it.
§§	By poor quality <u>cohort</u> study we mean one that failed to clearly define comparison groups and/or failed to measure exposures and outcomes in the same (preferably blinded), objective way in both exposed and non-exposed individuals and/or failed to identify or appropriately control known confounders and/or failed to carry out a sufficiently long and complete follow-up of patients. By poor quality <u>case-control</u> study we mean one that failed to clearly define comparison groups and/or failed to measure exposures and outcomes in the same (preferably blinded), objective way in both cases and controls and/or failed to identify or appropriately control known confounders.
§§§	Split-sample validation is achieved by collecting all the information in a single tranche, then artificially dividing this into "derivation" and "validation" samples.
††	An "Absolute SpPin" is a diagnostic finding whose <u>S</u> pecificity is so high that a <u>P</u> ositive result rules- <u>in</u> the diagnosis. An "Absolute SnNout" is a diagnostic finding whose <u>S</u> ensitivity is so high that a <u>N</u> egative result rules- <u>out</u> the diagnosis.
‡‡	Good, better, bad and worse refer to the comparisons between treatments in terms of their clinical risks and benefits.
†††	<u>Good</u> reference standards are independent of the test, and applied blindly or objectively to applied to all patients. <u>Poor</u> reference standards are haphazardly applied, but still independent of the test. Use of a non-independent reference standard (where the 'test' is included in the 'reference', or where the 'testing' affects the 'reference') implies a level 4 study.
††††	Better-value treatments are clearly as good but cheaper, or better at the same or reduced cost. Worse-value treatments are as good and more expensive, or worse and the equally or more expensive.
**	Validating studies test the quality of a specific diagnostic test, based on prior evidence. An exploratory study collects information and trawls the data (e.g. using a regression analysis) to find which factors are 'significant'.
***	By poor quality prognostic cohort study we mean one in which sampling was biased in favour of patients who already had the target outcome, or the measurement of outcomes was accomplished in <80% of study patients, or outcomes were determined in an unblinded, non-objective way, or there was no correction for confounding factors.
****	Good follow-up in a differential diagnosis study is >80%, with adequate time for alternative diagnoses to emerge (eg 1-6 months acute, 1 - 5 years chronic)

Grades of Recommendation

A	consistent level 1 studies
B	consistent level 2 or 3 studies or extrapolations from level 1 studies
C	level 4 studies or extrapolations from level 2 or 3 studies
D	level 5 evidence or troublingly inconsistent or inconclusive studies of any level

"Extrapolations" are where data is used in a situation which has potentially clinically important differences than the original study situation.