

Aus der Abteilung Diagnostische Radiologie
(Komm. Leiter: Prof. Dr. med. M. Knauth)
im Zentrum Radiologie
der Medizinischen Fakultät der Universität Göttingen

Studie zur klinischen Wertigkeit des CAD-Systems
in der digitalen Vollfeldmammographie
in Abhängigkeit von der Erfahrung des Befunders

INAUGURAL – DISSERTATION

zur Erlangung des Doktorgrades
der Medizinischen Fakultät der
Georg-August-Universität zu Göttingen

vorgelegt von

Besim Cetin **Angic**

aus
Böblingen

Göttingen 2010

Dekan: Prof. Dr. med. C. Frömmel

I. Berichterstatterin: PD Dr. med. S. Obenauer

II. Berichterstatter/in:

III. Berichterstatter/in:

Tag der mündlichen Prüfung:

Inhaltsverzeichnis

1. EINLEITUNG	1
2. DIAGNOSESTUDIEN	3
2.1 SENSITIVITÄT UND SPEZIFITÄT	3
2.2 ROC-ANALYSE UND AREA UNDER THE CURVE (AUC)	3
3. MATERIAL UND METHODE	6
3.1 DIGITALES MAMMOGRAPHIEGERÄT UND BILDDARSTELLUNG	6
3.2 COMPUTERASSISTIERTER DIAGNOSE (CAD)	7
4. DIAGNOSESTUDIE „MIT UND OHNE CAD“	9
4.1 STUDIENDESIGN	9
5. ERGEBNISSE	12
5.1 DESKRIPTIVE ERGEBNISSE.....	12
5.2 VERGLEICH DER AUC ANHAND DER ROC-ANALYSE IN ABHÄNGIGKEIT VON DER ERFAHRUNG DES BEFUNDERS, DER ANWENDUNG EINES CAD-SYSTEMS SOWIE VOM PATIENTENKOLLEKTIV	15
5.3 ZEITAUFWAND DER BEFUNDUNGEN IN ABHÄNGIGKEIT VON DER ERFAHRUNG DES BEFUNDERS, DER ANWENDUNG EINES CAD-SYSTEMS SOWIE VOM PATIENTENKOLLEKTIV	20
5.4 VERTEILUNG DER RICHTIG-POSITIVEN UND FALSCH-POSITIVEN MARKIERUNGEN IN ABHÄNGIGKEIT VOM JEWEILIGEN PATIENTENKOLLEKTIV	24
6. BILDERTEIL	27
7. DISKUSSION	35
8. ZUSAMMENFASSUNG	40
ANHANG	42
LITERATURVERZEICHNIS.....	42
TABELLENVERZEICHNIS.....	46
ABBILDUNGSVERZEICHNIS.....	47
ABKÜRZUNGSVERZEICHNIS.....	49
DANKSAGUNG	50
LEBENSLAUF	51

1. Einleitung

In Deutschland wurde im Jahr 2006 ein flächendeckendes Brustkrebs-Screeningsprogramm zur Bekämpfung des Brustkrebses, der nach wie vor häufigsten Krebserkrankung der Frauen, eingeführt. Das Mammographiescreening wird in Deutschland bislang sehr erfolgreich durchgeführt und stößt nach Rückmeldung der ersten Zentren auf reges Interesse der untersuchten Frauen (Diekmann S, Diekmann F 2008). In der diagnostischen Radiologie hat sich in den letzten Jahren in fast allen Bereichen die digitale Bildgebung etabliert. Besonders im Rahmen der Röntgenuntersuchung der weiblichen Brust, bei der sehr hohe Ansprüche an die Geräte- und Aufnahmetechnik gestellt werden, verspricht man sich Vorteile durch die digitale Vollfeldmammographie. Heutzutage stützt sich die Diagnostik des Mammakarzinoms auf den klinischen Befund, das Ergebnis der Mammographie- und der Ultraschalluntersuchungen. Als Standardaufnahmen der weiblichen Brust hat sich in diesem Zusammenhang die Schrägaufnahme (mediolateral oblique = MLO) in Kombination mit der kraniokaudalen Aufnahme (CC) international etabliert. Zur Qualitätssicherung der mammographischen Befundauswertung schreiben die Richtlinien der European Society of Mastology (EUSOMA) eine Doppelbefundung vor, wodurch die Sensitivität um ca. 15% gesteigert werden kann (Funovics et al. 2001). Dies ist jedoch zeitaufwendig, kostenintensiv und nur mit hohem personellem und logistischem Aufwand durchführbar. Eine mögliche Alternative zur ärztlichen Zweitbefundung stellen computerunterstützte Auswertesysteme (computerassistierte Diagnose, CAD) dar (Thurfjell et al. 1994). Die gegenwärtige Anwendungsoption dieser computerassistenten Detektionssysteme liegt in der Kennzeichnung von potentiell malignen Strukturen. Dies dient in erster Linie dazu, den Radiologen auf suspekte Befunde hinzuweisen. Der in diesem Zusammenhang positive Effekt der Anwendung von CAD-Systemen auf die Sensitivität wurde in vielen Studien bestätigt (Obenauer et al. 2005). Ein Problem bei der Erkennung und dem Ausschluss von suspekten Befunden stellt die häufige falsch-positive Markierung dar. Dies führt dazu, dass mit der Untersuchung eines CAD-Befundes oft ein hoher Zeitaufwand verbunden ist, da die falsch-positiven Markierungen den Radiologen von der Befundung ablenken. Der Zeitaufwand bei der Mammographiebefundung ist heute einer der wesentlichsten Faktoren im klinischen Alltag. Durch technische Weiterentwicklungen sind die modernen Workstations mit ihrer Software so programmiert und eingestellt, dass Minimalläsionen wie Mikrokalk und Verdichtungen erkannt und gleichzeitig markiert werden können. Mit Hilfe der computerassistenten Diagnosestellung ist man in der Lage, die Treffsicherheit zu erhöhen. Diese technischen Weiterentwicklungen sind derzeit noch nicht im Stande, den menschlichen

Befunder zu ersetzen, sondern helfen dabei, eine schnellere und sicherere Diagnose zu stellen. Mikrokalk entsteht aus Überresten abgestorbener Zellen und kann einen Hinweis auf ein malignes Geschehen geben. Dabei muss man Anzahl, Form, Kontrast und geometrische Anordnung betrachten, da diese mit dem Risiko der Malignität korrelieren. Dennoch kommen solche Mikrokalzifikationen auch bei gutartigen Läsionen wie Zysten oder Mastopathien zur Darstellung. Oft ist die Biopsie des entsprechenden Gewebes die einzige Möglichkeit zur Feststellung, ob diese Veränderung in der Brust gut- oder bösartig ist (Giger 1993). Um Verdichtungsherde (Herdbefunde) zu erkennen und anzuzeigen, gelten besondere Merkmale wie Konturschärfe, Form, Größe, Dichte und Verteilung der Verdichtungen. Das CAD-System Image Checker R2 markiert suspekte Verdichtungen, die radiär auslaufende Strahlen aufweisen und eine bestimmte Größe besitzen. So müssen die Verdichtungen mindestens größer als 6mm sein und dürfen einen maximalen Durchmesser von 3,2cm besitzen. Im Rahmen dieser Arbeit wird eine prospektive zwei-faktorielle Diagnosestudie mit drei unterschiedlich erfahrenen Befundern durchgeführt, in welcher jeder Befunder mit und ohne CAD die Aufnahmen bewertet. Ziel ist es, die Sensitivität, Spezifität, die diagnostische Güte und die Befundungszeiten der computerassistierten Diagnose in Abhängigkeit der Untersuchererfahrung und der Tumormanifestation zu evaluieren. Um die Sensitivität und die Spezifität der drei Befunder mit und ohne das CAD-System zu bewerten, wurde eine ROC-Analyse mit der entsprechenden Fläche unter der Kurve AUC (AUC = area under the curve) für jeden einzelnen Befunder erstellt und mit Hilfe von speziellen Varianzanalysen untersucht, ob sich die (erwarteten) AUC's im Bezug auf die beiden Methoden, die drei Befunder und die Wechselwirkung zwischen Befunder und Methode, statistisch unterscheiden. Hierzu wird die von Werner (2006) vorgestellte Methodik angewendet.

2. Diagnosestudien

Diagnosestudien dienen der Untersuchung, Evaluierung und die Bewertung diagnostischer Verfahren in der Medizin, wobei die diagnostische Güte Gegenstand des Interesses ist. Bei Diagnosestudien handelt es sich aus zweierlei Gründen um eine besonders interessante und anspruchsvolle Art klinischer Studien. Das verwundert zunächst, geht es doch letztendlich nur um die Auswertung einer Vierfeldertafel, in der zwei Zustände des Patienten (krank D+, nicht krank D-) zwei Ergebnissen des diagnostischen Tests gegenübergestellt sind (test positiv T+, test negativ T-). Im Folgenden werden kurz wichtige Gütemaße der Diagnostik – Sensitivität, Spezifität und die AUC kurz vorgestellt, mit welchen die durchgeführte Diagnosestudie ausgewertet wird.

2.1 Sensitivität und Spezifität

CAD-Markierungen, die mit dem mammographischen Befund der histologisch gesicherten Karzinome übereinstimmten, wurden als „richtig-positive Treffer“ gedeutet. Im Gegensatz dazu wurden CAD-Markierungen in anderen nicht malignen Parenchymabschnitten als „falsch-positive“ gewertet und fehlende Markierungen, bei welchen das System einen histologisch gesicherten Befund übersehen hat, als „falsch-negativer Treffer“ angesehen. Die Anzahl der „Treffer“ steht stellvertretend für die Summe der richtig-positiven Markierungen oder die Sensitivität in Prozent. In der Statistik steht die Sensitivität (bzw. die richtig-positive Rate) eines statistischen Tests für die Wahrscheinlichkeit, dass ein positives Ergebnis auch als ein solches erkannt wird (Werner 1992):

P (positiv erkannt / richtig-positiv).

Im Gegensatz dazu versteht man unter der Spezifität die Wahrscheinlichkeit, dass ein negatives Ergebnis auch als ein solches erkannt wird (Werner 1992):

P (negativ erkannt / richtig-negativ).

2.2 ROC-Analyse und Area under the curve (AUC)

Die ROC-Kurve („Receiver Operating Characteristic“), gibt einen Überblick über Sensitivität und Spezifität eines diagnostischen Tests. In ihr werden für verschiedene Cut-Off-Werte Sensitivität und Spezifität gegeneinander aufgetragen. Ein Maß für die Güte des Tests ist die Fläche unter der ROC-Kurve (AUC = area under the curve). Die Fläche kann Werte zwischen 0,5 und 1 annehmen, wobei ein höherer Wert die bessere Güte anzeigt. Anhand der ROC-Kurve ist es möglich, die Abhängigkeit der Effizienz von der Fehlerrate darzustellen. Bei der

ROC-Analyse ermittelt man für jeden möglichen Cut-off Punkt die daraus resultierenden relativen Häufigkeitsverteilungen und kann dadurch die Sensitivität sowie Spezifität errechnen. Die Ergebnisdarstellung erfolgt anhand eines Diagramms, in dem die Ordinate die „Sensitivität“ und die Abszisse die 1-„Spezifität“ darstellen. Um nun die Güte eines Diagnoseverfahrens zu bestimmen, besteht die Möglichkeit, Sensitivität und Spezifität für einen festen Wert c zu betrachten. Wird der Schwellenwert hierbei allerdings falsch festgesetzt, so kann das Urteil über ein sonst gutes diagnostisches Verfahren negativ ausfallen. Daher wird als Index für die Treffsicherheit eines Diagnoseverfahrens die Fläche unter der ROC-Kurve (AUC = area under the curve) verwendet, denn bei diesem Gütemaß werden Sensitivität und Spezifität für alle möglichen Werte c berücksichtigt. Dieser Index hat außerdem eine anschauliche Interpretation: Er gibt die Wahrscheinlichkeit an, dass bei je einem zufällig gewählten kranken und gesunden Patienten $X_{\text{krank}} > X_{\text{gesund}}$ gilt, d.h. dass bei dem kranken Patienten ein größerer Wert als bei dem gesunden gemessen wird. Somit gibt die AUC Auskunft über die Trennschärfe (Accuracy) eines Diagnoseverfahrens über den gesamten Wertebereich. Berechnet wird die AUC dabei durch: $AUC = P(X_{\text{gesund}} < X_{\text{krank}}) + 0.5 * P(X_{\text{gesund}} = X_{\text{krank}})$.

Mögliche ROC-Kurven sind in der Abbildung 1 dargestellt, die Werner (2006) entnommen wurde. Die Verteilung der kranken ist jeweils rot, die der gesund blau dargestellt.

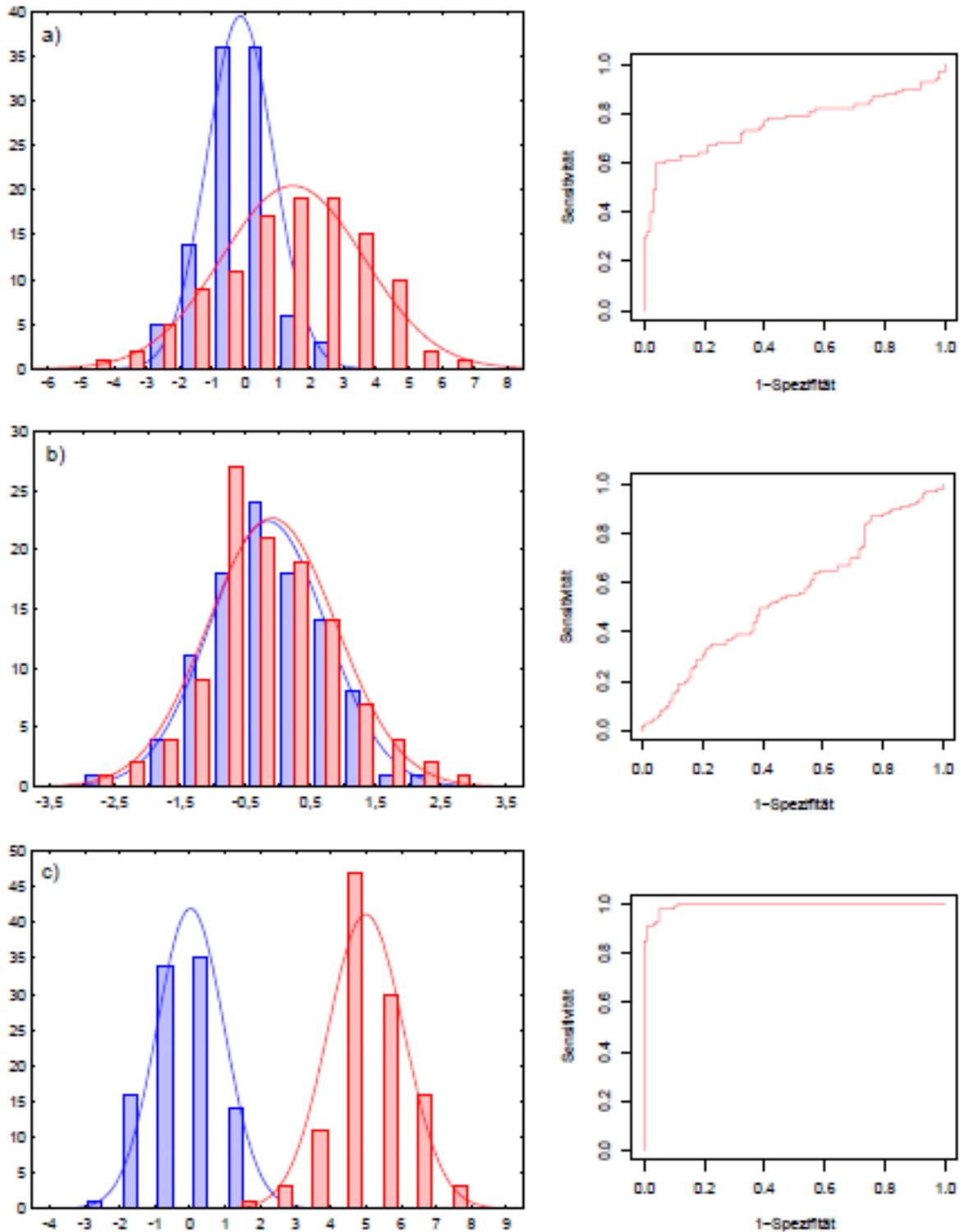


Abbildung 1: Empirische Dichten und ROC-Kurven: kranke Patientinnen (rote Balken) und gesunde (blaue Balken). (a) Standardbeispiel, (b) Extremfall 1: nicht besser als der Zufall, (c) Extremfall 2: perfekte Trennung.

3. Material und Methode

3.1 Digitales Mammographiegerät und Bilddarstellung

Die Untersuchungen für diese Arbeit wurden an einem Gerät zur digitalen Vollfeldmammographie (Senographe 2000D, Firma GE, Medical Systems, Buc, Frankreich; Abbildung 2) durchgeführt. Die Systemausstattung des Senographe 2000 D besteht aus einer Bimetallanode, aus Rhodium (Rh) und Molybdän (Mo). Von diesen beiden Anoden wird immer nur jeweils ein Anodenmaterial angeregt. Als Filter kommt ebenfalls Molybdän (0,03mm) und Rhodium (0,025mm) zum Einsatz. Die Röhrenspannung beträgt 22-49kV. Das digitale System verwendet als Detektormaterial eine Kombination von amorphem Silizium und einem Cäsiumjodid-Szintillator. Bei einer aktiven Fläche von 19cm x 23cm beträgt die Pixelgröße 100µm. Durch diese Kombination erhält man eine Detektormatrix von 1.900 x 2.300 Bildpunkten und ein Ortsauflösung von maximal 5 Linienpaaren pro Millimeter (Lp/mm) bei Übersichtsaufnahmen. Bei dem Senographe 2000 D erreicht der Detektor eine relativ hohe Quantenausbeute mit einer Dosisquanteneffizienz (DQE) von 55% bei 0 Lp/mm und besitzt einen großen Dynamikbereich von > 10.000: 1. Die digital angefertigten Aufnahmen können an zwei 2kx2,5k Monitoren angeschaut und befundet werden oder auch als Hardcopy über einen hochauflösenden Laserdrucker der Firma AGFA Leverkusen (Typ Scopix LR 5200) ausgedruckt und archiviert werden. Der Laserdrucker erreicht eine Pixelgröße von 40µm (Obenauer et al. 2003).



Abbildung 2: Senographe 2000D, Firma GE, Medical Systems, Buc, Frankreich.

3.2 Computerassistierte Diagnose (CAD)

In dieser Studie kam das CAD-System Image Checker der Firma R2 Technology, Los Altos, Kalifornien, USA, zum Einsatz. Es handelt sich hierbei um ein System, das in der Radiologischen Abteilung der Universität Göttingen in Kombination mit dem volldigitalen Mammographiegerät Senographe 2000 D der Firma GE zum Einsatz kommt. Die Software des CAD-Systems Image Checker R2 basiert auf speziellen Algorithmen, die durch Segmentation der Mammogramme in der Lage sind, Läsionen wie Herdbefunde und Mikroverkalkungen zu erkennen und in wenigen Sekunden darzustellen. Das CAD-System zeigt Mikroverkalkungen an, wenn mindestens drei oder mehr Kalkpartikel auf einer Fläche von $0,5\text{cm}^2$ erscheinen. Herdbefunde müssen größer als 6mm sein und einen maximalen Durchmesser von 3,2cm besitzen, um markiert zu werden.

Die Darstellung des CAD-Systems Image Checker R2 erfolgt an einer Mammographie-Workstation, an der zwei separate, hochauflösende Monitore angeschlossen sind (Abbildung 3). Die Markierungen werden jeweils als Dreieck für Mikroverkalkungen (▲) und als Stern für Herdbefunde (*) gesetzt. Es werden maximal drei Mikrokalzifikationen und zwei Herdbefunde pro Bild vom CAD-System Image Checker R2 angezeigt.



Abbildung 3: Mammographie Workstation mit zwei hochauflösenden Monitoren (2000×2500 Pixel) für die Befundung. Zusätzlich abgebildet ist ein Computer zur Bedienung des radiologischen Informationssystems.

4. Diagnosestudie „mit und ohne CAD“

4.1 Studiendesign

Der Aufbau der Studie erfolgte durch Rekrutierung eines repräsentativen Patientenkollektivs. Dieses bestand aus der Früherkennungsgruppe (n = 103), deren Frauen sich rein zur Vorsorge vorstellten und keine kranke Brust hatten. Frauen der Benignengruppe (n=102) hatten einen gutartigen Befund in einer- oder beiden Brüsten. Die dritte Gruppe stellt die Malignengruppe (n=98) dar. Diese Patientinnen trugen einen bösartigen Befund in einer oder beiden Brüsten. Die Diagnosen sind histologisch gesichert. In dieser Diagnosestudie gehen Beobachtungen der Früherkennungsgruppe als *gesunde (d.h. Goldstandard = gesund)* Beobachtungen, die der malignen Frauen als *kranke oder gesunde (d.h. Goldstandard = krank oder gesund)* Beobachtungen in die Analyse ein. Die Ebene des Goldstandards (gesund / krank) ist immer die Brust selbst. Von allen Patientinnen wurde eine digitale Mammographie beider Brüste in den Standardebenen kraniokaudal (CC), sowie mediolateral oblique (MLO) erstellt. Alle Aufnahmen wurden randomisiert und in zufälliger Reihenfolge drei Befundern zur Diagnose vorgelegt, die jeweils mit und ohne Unterstützung des CAD-Systems durchgeführt wurde. Die Aufnahmen wurden nach dem gängigen Standard der Breast Imaging Reporting and Data System (BI-RADS) des American College of Radiology (ACR) I bis V klassifiziert (Tabelle 1).

Tabelle 1: BI-RADS Klassifikation (D' Orsi et al. 1998)

BI-RADS	Bewertung	Procedere	Karzinomrisiko
0	Inkomplette Diagnostik	Komplette Diagnostik anstreben	Unklar
I	Kein Befund	Keine weiteren Kontrollen	0 %
II	Sicher benigner Befund	Keine weiteren Kontrollen	0 %
III	Wahrscheinlich benigner Befund	Kontrolle in 6 Monaten	1-3 %
IV	Wahrscheinlich maligner Befund	Biopsie (Histologie)	3-90 %
V	Malignomtypischer Befund	Adäquate Therapie	Ca. 90 %

Die drei Befunder haben einen unterschiedlichen Ausbildungsstand in der Mammographie. Der Befunder 1 hat als Assistenzarzt eine mehrjährige Berufserfahrung, wohingegen der Befunder 2 als Student lediglich eine kurze Einarbeitungszeit besitzt. Der Befunder 3 hat als Facharzt mit Spezialkenntnissen die größte Erfahrung.

Jede Aufnahme wurde von jedem Befunder mit jeder Methode (mit und ohne CAD-System) befundet. Da jedes Bild nicht nur einmal, sondern wiederholt gemessen wurde, sind die Daten nicht unabhängig, sondern abhängig. Da weiterhin beide Brüste der Patientinnen begutachtet wurden, kann es vorkommen, dass manche Patientinnen sowohl gesunde als auch kranke Beobachtungen liefern, andere Patientinnen hingegen liefern nur gesunde Beobachtungen. Solche speziellen Datenstrukturen werden auch *clustered-data* genannt.

Eine anschauliche Darstellung der Datenvorlage bringt die folgende Abbildung 4 die der Arbeit von Werner (2006) entnommen wurde:

	Methode 1						Methode 2					
	Reader 1		Reader 2		Reader 3		Reader 1		Reader 2		Reader 3	
Sub.	gesund	krank	gesund	krank	gesund	krank	gesund	krank	gesund	krank	gesund	krank
1	xxx	xx	xxx	xx	xxx	xx	xxx	xx	xxx	xx	xxx	xx
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n_c	xx	xxxx	xx	xxxx	xx	xxxx	xx	xxxx	xx	xxxx	xx	xxxx
1	xxx		xxx		xxx		xxx		xxx		xxx	
⋮	⋮		⋮		⋮		⋮		⋮		⋮	
$n_0 - n_c$	xx		xx		xx		xx		xx		xx	
1		xxxx		xxxx		xxxx		xxxx		xxxx		xxxx
⋮		⋮		⋮		⋮		⋮		⋮		⋮
$n_1 - n_c$		xxx		xxx		xxx		xxx		xxx		xxx

Abbildung 4: Struktur von clustered-data, x steht für eine Beobachtung (Werner 2006).

Die Befunder und die beiden Diagnosemethoden werden hier als feste Einflussfaktoren betrachtet, wodurch wir eine 2-faktorielle Diagnosestudie mit so genannten clustered-data erhalten. Die Faktoren Methode, Befunder und die Wechselwirkung Methode * Befunder sollen in der statistischen Analyse untersucht werden. Für jede Befunder * Methoden Kombination wird eine separate AUC geschätzt so dass mit Hilfe einer Varianzanalyse folgende Hypothesen statistisch überprüft werden sollen:

1. H_0^M : Die diagnostische Güte beider Methoden sind gleich (d.h. die beiden Methoden-AUC's, jeweils gemittelt über die Befunder, sind gleich).

2. H_0^R : Die diagnostische Güte der drei Befunder ist gleich (d.h. die drei AUC's der Befunder, jeweils gemittelt über die beiden Methoden, sind gleich).
3. H_0^{MR} : Die Diagnose der Befunder ist mit den beiden Methoden homogen.

In der Arbeit von Werner (2006) werden statistische Verfahren (Varianzanalysen für clustered-data) vorgestellt, die hier angewendet werden können.

5. Ergebnisse

5.1 Deskriptive Ergebnisse

Die Früherkennungsgruppe setzte sich aus 103 Patientinnen zusammen, die sich zu regelmäßigen Vorsorgeuntersuchungen vorstellten und jederzeit, auch im Verlauf mit BI-RADS 1 oder 2 klassifiziert wurden. Die Altersverteilung der Früherkennungsgruppe ist in Abbildung 5 dargestellt.

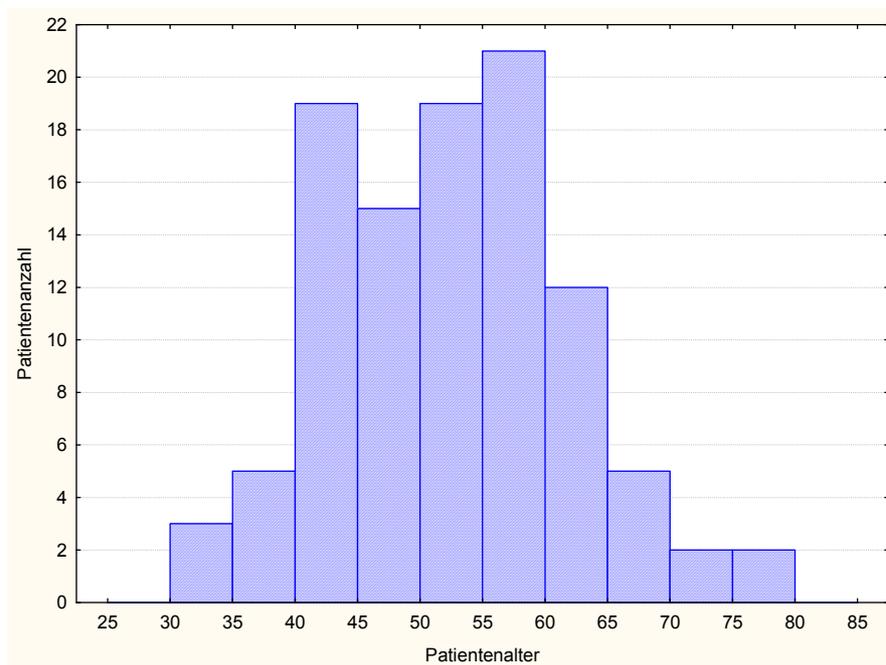


Abbildung 5: Altersverteilung der Früherkennungsgruppe.

Die Benignengruppe bestand aus 102 Patientinnen, die sich aufgrund von symptomatischen Brustbeschwerden zur Abklärung einer digitalen Vollfeldmammographie mit anschließender Probebiopsie unterzogen haben. Die Altersverteilung sowie Histologie sind in Abbildung 6 und 7 aufgeführt.

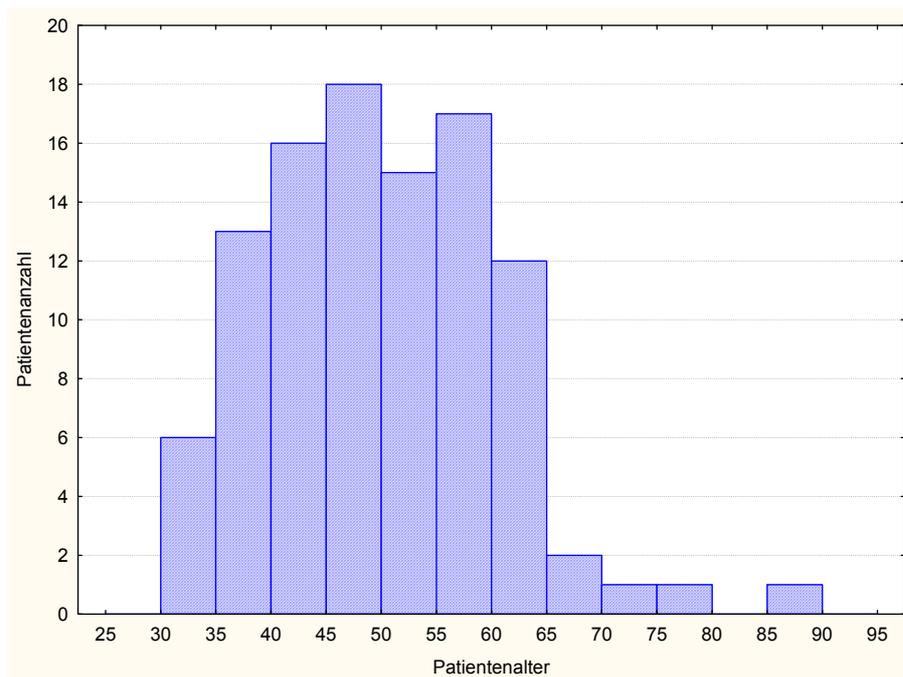


Abbildung 6: Altersverteilung der Benignengruppe.

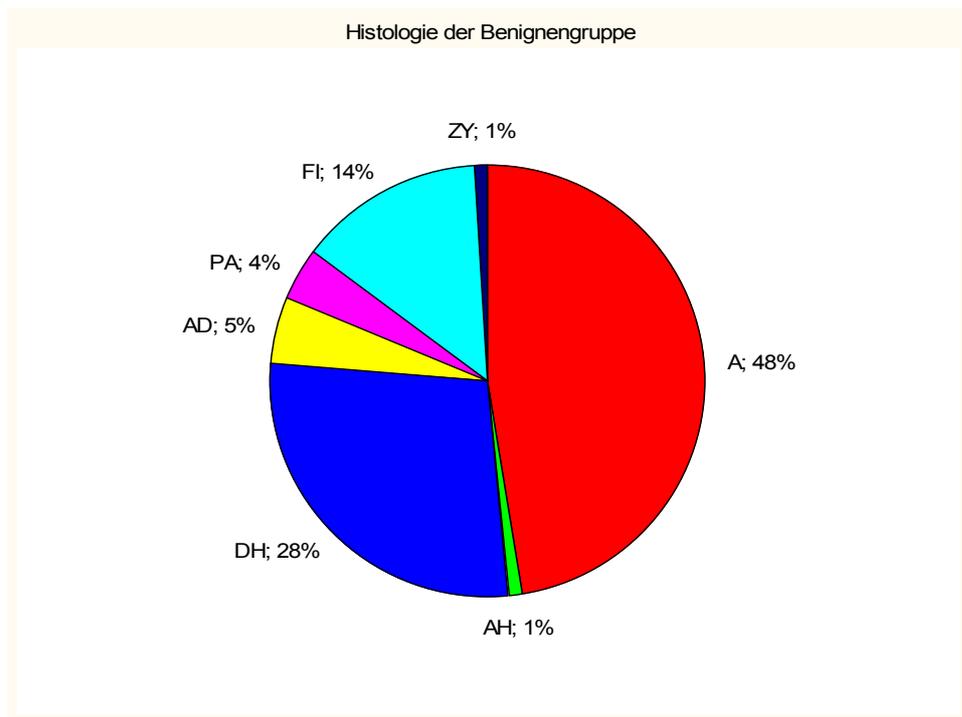


Abbildung 7: Histologie der Benignengruppe (A= Adenom; DH= dukta Hyperplasie; FI = Fibroadenom; AD = Adenose, PA = Papillom; AH = atypische dukta Hyperplasie; Z = Zyste)

Die Malignengruppe bestand aus 98 Patientinnen mit histologisch gesichertem Mammakarzinom (BI-RADS 4 und 5). Bei den erkrankten Patientinnen wurde präoperativ

eine digitale Vollfeldmammographie durchgeführt. Die Altersverteilung der Malignengruppe sowie Histologie sind in Abbildung 8 und 9 dargestellt.

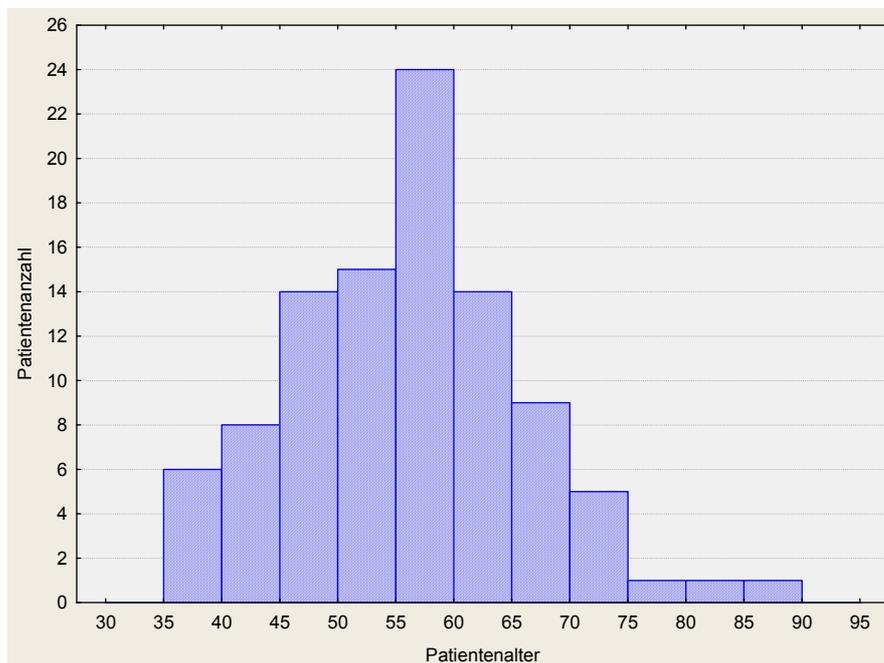


Abbildung 8: Altersverteilung der Malignengruppe.

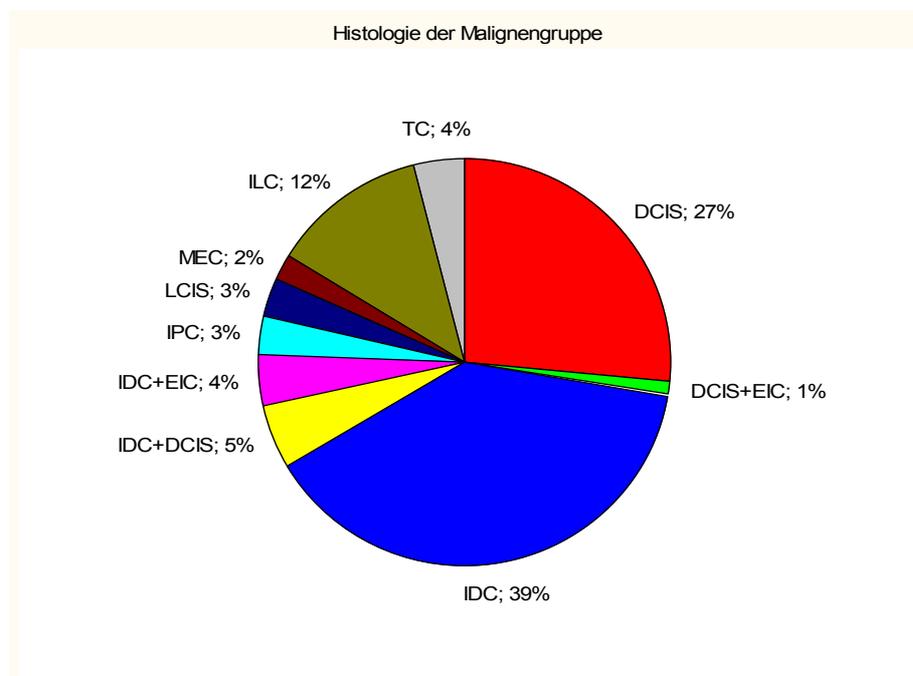


Abbildung 9: Histologie der Malignengruppe.

(IDC = invasiv-duktales Karzinom; DCIS = duktales Carcinoma in situ; ILC = invasiv- lobuläres Karzinom; IDC+DCIS = invasiv-duktales Karzinom + duktales Carcinoma in situ; IDC+EIC = invasiv-duktales Karzinom + extensive intraduktale Komponente ; DCIS+EIC = duktales Carcinoma in situ + extensive intraduktale Komponente; MEC = medulläres extensives Karzinom; LCIS = lobuläres Carcinoma in situ; IPC = invasiv papilläres Karzinom; TC = tubuläres Karzinom)

5.2 Vergleich der AUC anhand der ROC-Analyse in Abhängigkeit von der Erfahrung des Befunders, der Anwendung eines CAD-Systems sowie vom Patientenkollektiv

Analyse der Sensitivitätsveränderung durch Anwendung des CAD-Systems in der Malignengruppe und der Früherkennungsgruppe.

In der folgenden Untersuchung wurden 201 Patienten analysiert. Dieses Kollektiv setzte sich aus der Malignen- (n=98) und der Früherkennungsgruppe (n=103) zusammen. Abbildung 10 zeigt die empirischen ROC-Kurven aller drei Befunder mit und ohne Unterstützung des CAD-Systems. Die Mehrfachmessungen wurden für die grafische Veranschaulichung jeweils gemittelt. Dadurch dass BI-RADS ordinale Daten sind, verlaufen die empirischen ROC-Kurven nicht stetig, sondern punktweise. Jeder Punkt der ROC-Kurve ist ein Cut-Off Punkt durch die BI-RADS Graduierung. Zur leichteren Orientierung wurde die 1. Winkelhalbierende mit in die Grafiken aufgenommen.

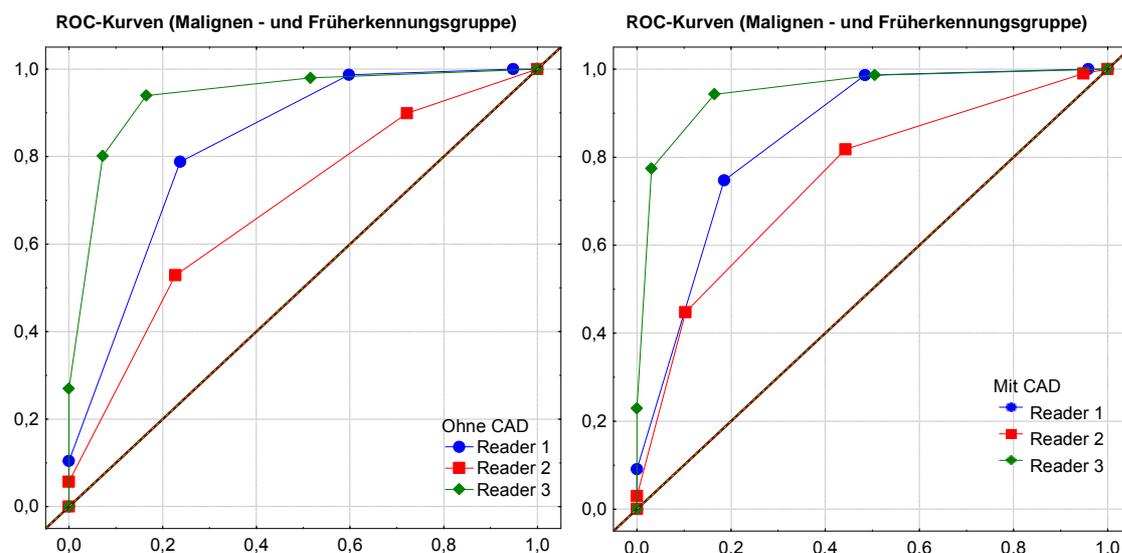


Abbildung 10: ROC-Kurven aller drei Befunder für die Malignen- und Früherkennungsgruppe ohne (links) und mit (rechts) Gebrauch des CAD-Systems.

Die Abbildung zeigt, dass der Student mit und ohne CAD den flachsten Verlauf der Kurve und infolgedessen die geringste AUC besitzt. Mit der Unterstützung des CAD-Systems ist die ROC-Kurve steiler. Der Assistenzarzt erreicht mit und ohne den Einfluss von CAD einen steileren Anstieg des Diagramms als der Student. Unter der Zuhilfenahme des CAD-Systems ist die Kurve auch steiler als ohne CAD-Anwendung. Dies zeigt, dass der Assistenzarzt eine höhere AUC als der Student erzielt. Der steilste Anstieg der ROC-Kurve und infolgedessen die höchste AUC hat der Facharzt mit und ohne den CAD-Einsatz. Die berechneten AUC's

für jeden Befunder und für jede Methode, sowie der Obere als auch der Untere Bereich der Konfidenzintervalle sind in Tabelle 2 abgebildet.

Tabelle 2: AUC aller drei Befunder in Abhängigkeit der Fläche unter der Kurve (AUC= area under the curve) für die Malignen - und Früherkennungsgruppe mit und ohne Anwendung des CAD-Systems

Methode	Befunder	AUC	95% Untergrenze	95% Obergrenze
ohne CAD	Assistenzarzt	0.809	0.767	0.846
	Student	0.649	0.606	0.690
	Facharzt	0.894	0.847	0.928
mit CAD	Assistenzarzt	0.834	0.785	0.873
	Student	0.712	0.668	0.752
	Facharzt	0.911	0.876	0.936

Tabelle 2 zeigt, dass alle drei Befunder einen Anstieg der AUC haben, wenn sie das CAD-System einsetzen. Den höchsten Nutzen unter dem Einsatz des CAD-Systems hat der Student (10% Zunahme), gefolgt vom Assistenzarzt (3%) und schließlich vom Facharzt (2%).

Um auszuwerten, ob es signifikante Unterschiede in den diagnostischen Güten in Abhängigkeit der Erfahrung des Befunders, der Untersuchungsmethode mit und ohne Einsatz des CAD-Systems und der Interaktion Befunder und CAD-System gibt, werden die von Werner (2006) entwickelten Verfahren angewendet und die oben formulierten Hypothesen statistisch getestet.

Tabelle 3: Unterschiede in der AUC für die Malignengruppe und Früherkennungsgruppe bezogen auf die Befunder, Methode und die Interaktion Befunder und Methode

Hypothese	Frage	p-Wert	Interpretation
Methode: H_0^M	Unterscheiden sich die zwei Methoden?	0.0015	signifikant
Befunder: H_0^R	Unterscheiden sich die Befunder?	0.0001	signifikant
Methode*Befunder: H_0^{MR}	Ist die Diagnose der Befunder homogen?	0.1293	nicht signifikant

In Tabelle 3 werden die entsprechenden AUC's statistisch ermittelt und verglichen. Wie man aus der Tabelle entnehmen kann unterscheiden sich die Methode und die Befunder signifikant voneinander. Die Interaktion Methode*Befunder zeigt jedoch keinen signifikanten Unterschied. Im Folgenden wird eine Sensitivitätsanalyse durchgeführt, bei welcher nur die diagnostische Güte des CAD-Systems in der malignen Gruppe untersucht wird. Dieses bedeutet, dass hier das Patientenkollektiv nur aus den Frauen besteht, die sowohl gesunde als auch kranke Beobachtungen liefern. In Anlehnung an den vorigen Abschnitt, sieht die Datenvorlage folgendermaßen aus (Abbildung 11).

	Methode 1						Methode 2					
	Reader 1		Reader 2		Reader 3		Reader 1		Reader 2		Reader 3	
Sub.	gesund	krank	gesund	krank	gesund	krank	gesund	krank	gesund	krank	gesund	krank
1	xxx	xx	xxx	xx	xxx	xx	xxx	xx	xxx	xx	xxx	xx
.
.
$n_{i\cdot}$	xx	xxxx	xx	xxxx	xx	xxxx	xx	xxxx	xx	xxxx	xx	xxxx

Abbildung 11: Struktur von clustered-data, x steht für eine Beobachtungen.

Sensitivitätsanalyse mit und ohne Anwendung des CAD-Systems in der Malignengruppe auf der Grundlage der ROC-Kurven.

In dieser statistischen Analyse wurden 98 Patienten mit histologisch nachgewiesenem Brustkrebs untersucht. Die Abbildung 12 zeigt die empirischen ROC-Kurven der drei Befunder mit und ohne Unterstützung des CAD-Systems für die Malignengruppe. Die Abbildung zeigt, dass der Student mit und ohne das CAD-System die flachste Zunahme der Kurve hat. Ohne den Einfluss des CAD-Systems ist es eine fast lineare Steigung und mit CAD verläuft das Diagramm steiler. Der Student besitzt die flachste AUC. Der Assistenzarzt erreicht mit und ohne den Einfluss des CAD-Systems einen steileren Anstieg der AUC als der Student. Unter Verwendung des CAD-Systems verläuft die AUC ebenfalls steiler als ohne Anwendung des CAD-Systems. Dies zeigt, dass der Assistenzarzt eine höhere AUC besitzt als der Student. Der Facharzt erreicht den steilsten Anstieg in der AUC unabhängig ob mit und ohne Anwendung des CAD-Systems. Infolgedessen hat er die höchste Treffgenauigkeit als alle anderen Befunder. Sogar unter Berücksichtigung der Erfahrung des Facharztes verläuft das Diagramm durch Unterstützung des CAD-Systems steiler als ohne Anwendung des CAD-Systems.

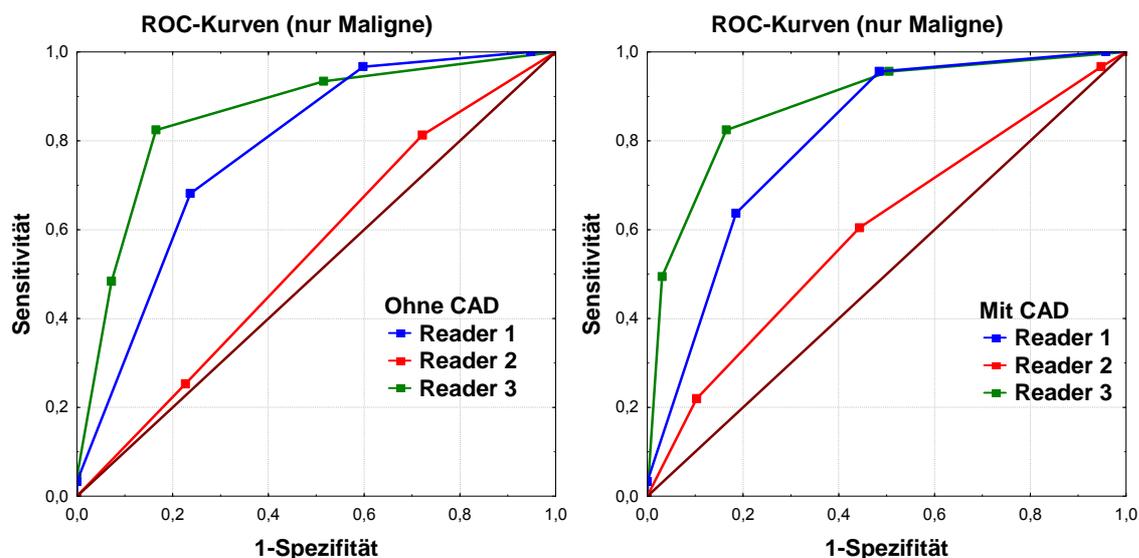


Abbildung 12: ROC-Kurven aller drei Befunder für die Malignengruppe ohne (links) und mit (rechts) Gebrauch des CAD-Systems.

Um den Einfluss auf die Treffgenauigkeiten der drei Befunder mit und ohne das CAD-System auszuwerten, wurde die AUC gemessen. Die berechnete AUC für jeden Befunder und für jede Methode, sowie die obere und untere Grenze der Konfidenzintervalle wird in Tabelle 4 gezeigt.

Tabelle 4: AUC aller drei Befunder in Abhängigkeit der Fläche unter der Kurve für die Malignengruppe mit und ohne Anwendung des CAD-Systems

Methode	Befunder	AUC	95%- Untergrenze	95%- Obergrenze
ohne CAD	Assistenzarzt	0.778	0.721	0.827
	Student	0.545	0.485	0.604
	Facharzt	0.856	0.781	0.909
mit CAD	Assistenzarzt	0.805	0.744	0.855
	Student	0.599	0.531	0.664
	Facharzt	0.883	0.820	0.926

Tabelle 4 zeigt, dass alle drei Befunder einen Anstieg in ihrer Treffgenauigkeit haben, wenn sie das CAD-System verwenden. Den höchsten Nutzen hat der Student (10% Zunahme),

gefolgt vom Assistenzarzt (4%) und schließlich gefolgt vom Facharzt (3%). Um auszuwerten, ob es signifikante Sensitivitätsunterschiede in Abhängigkeit von der Erfahrung des Befunders, Untersuchungsmethoden mit und ohne das CAD-System gibt, wurden die entsprechenden p-Werte bestimmt. Tabelle 5 stellt die p-Werte und die Interpretation der statistischen Signifikanz für die Malignengruppe dar.

Tabelle 5: Die unterschiedliche Treffgenauigkeit für die Malignengruppe und die Früherkennungsgruppe bezogen auf Befunder und Methode und Interaktion Befunder und Methode

Hypothese	Frage	p-Wert	Interpretation
Methode	Unterscheiden sich die zwei Methoden?	0.0157	signifikant
Befunder	Unterscheiden sich die Befunder?	0.0001	signifikant
Methode*Befunder	Ist der Verlauf homogen?	0.5958	nicht signifikant

5.3 Zeitaufwand der Befundungen in Abhängigkeit von der Erfahrung des Befunders, der Anwendung eines CAD-Systems sowie vom Patientenkollektiv

Die Tabelle 6 stellt die Mittelwerte und 95% Konfidenzintervalle der Befundungszeiten dar.

Tabelle 6: Abhängigkeit der Befundungszeiten vom jeweiligen Ausbildungsstand der Befunder

Gruppe	CAD	Befunder	N	Mean	95%- Unter- grenze	95%- Ober- grenze
Benignengruppe	Ohne	Assistenzarzt	102	26.53	24.72	28.34
		Student	102	35.19	33.89	36.48
		Facharzt	102	11.12	9.909	12.33
	Mit	Assistenzarzt	102	28.24	26.76	29.71
		Student	102	39.71	38.20	41.21
		Facharzt	102	10.93	9.735	12.13
Früherkennungs- gruppe	Ohne	Assistenzarzt	103	22.60	20.87	24.33
		Student	103	30.10	28.77	31.42
		Facharzt	103	7.466	6.528	8.404
	Mit	Assistenzarzt	103	23.56	22.00	25.12
		Student	103	32.12	30.72	33.52
		Facharzt	103	6.796	6.112	7.480
Malignengruppe	Ohne	Assistenzarzt	98	27.54	25.85	29.23
		Student	98	37.18	35.70	38.66
		Facharzt	98	11.37	9.624	13.12
	Mit	Assistenzarzt	98	30.43	28.60	32.26
		Student	98	43.00	41.59	44.41
		Facharzt	98	11.32	9.773	12.87

Eine graphische Darstellung der unterschiedlichen Befundungszeiten liefert die Abbildung 13. Zusätzlich werden der Median sowie die 25-75% Quantile und die Non-Outlier-Range der Verteilungen dargestellt. Die Abbildung zeigt eindeutig eine Abhängigkeit von der jeweiligen Erfahrung des Befunders und Patientengruppe, sowie einen Zeitunterschied bei der Untersuchung mit und ohne den Einsatz des CAD-Systems. Der Facharzt benötigt für die Befundung unabhängig vom CAD-Einsatz erwartungsgemäß die geringste Zeit, gefolgt vom Assistenzarzt und vom Studenten. In allen 3 Patientengruppen verlängert die CAD-

Anwendung die Befundungszeit des Studenten und des Assistenzarztes. Der Facharzt hingegen benötigt mit und ohne CAD-Technik im Median die gleiche Zeit.

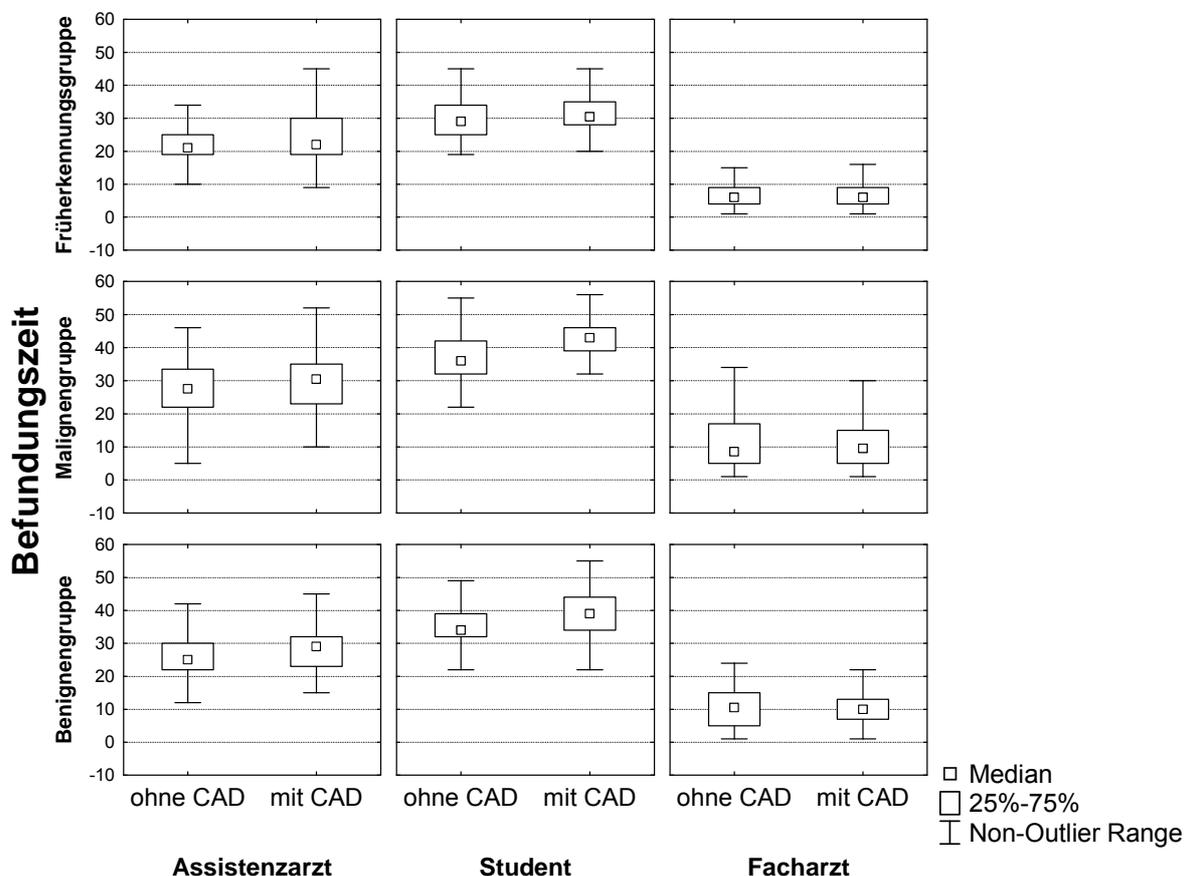


Abbildung 13: Box-Plots der Befundungszeiten in Abhängigkeit von der Patientengruppe, vom Befunder sowie von der Anwendung des CAD-Systems.

Mit Hilfe eines gemischten Modells soll nun statistisch untersucht werden, ob sich

- 1.) die Befundungszeiten zwischen der Diagnose mit und ohne CAD und
- 2.) die Befundungszeiten der drei Befundern unterscheiden und ob
- 3.) die Befundungszeiten zwischen den drei Befundern in den beiden Methoden homogen sind.

Die folgenden Auswertungen zeigen getrennt für die einzelnen 3 Gruppen die p-Werte für die eben formulierten Hypothesen (Tabelle 7 bis 9).

Tabelle 7: Unterschiede in den Befundungszeiten in der Früherkennungsgruppe in Abhängigkeit vom Befunder sowie von der Anwendung des CAD-Systems

Hypothese	P-Wert	Interpretation
Befundungszeiten der Methoden	0.1334	nicht signifikant
Befundungszeiten der Befunder	< 0.0001	signifikant
Methode*Befunder	0.1001	nicht signifikant

Bei der Früherkennungsgruppe gab es keine signifikanten unterschiedlichen Befundungszeiten bezüglich der Anwendung von CAD ($p= 0.1343$). Bezüglich der Erfahrung des Befunders gibt es allerdings einen deutlichen signifikanten Unterschied ($p= 0.0001$). Der Einsatz der CAD-Technik korreliert signifikant mit der jeweiligen Befundungserfahrung, allerdings ergeben sich keine signifikanten Wechselwirkungen zwischen den Erfahrungswerten der drei Befunder und dem Einsatz des CAD-Systems (Methode). Im Vergleich dazu liefert der p-Wert bei der Benignengruppe (Tabelle 8) und bei der Malignengruppe (Tabelle 9) signifikante Wechselwirkungen zwischen den Erfahrungswerten der drei Befunder und dem Einsatz der CAD-Technik.

Tabelle 8: Unterschiede in den Befundungszeiten in der Benignengruppe in Abhängigkeit vom Befunder sowie von der Anwendung des CAD-Systems

Hypothese	P-Wert	Interpretation
Befundungszeiten der Methoden	0.0003	signifikant
Befundungszeiten der Befunder	< 0.0001	signifikant
Befunder*Methode	0.0023	signifikant

Tabelle 9: Unterschiede in den Befundungszeiten in der Malignengruppe in Abhängigkeit vom Befunder sowie von der Anwendung des CAD-Systems

Hypothese	P-Wert	Interpretation
Befundungszeiten der Methoden	< 0.001	signifikant
Befundungszeiten der Befunder	< 0.001	signifikant
Befunder*Methode	0.0015	signifikant

Die Tabelle 6 bis Tabelle 9 und Abbildung 13 zeigen deutlich, dass die Befundungszeit unter Anwendung der CAD-Auswertung in allen 3 Patientengruppen beim Assistenzarzt und beim Studenten signifikant erhöht ist. Beim Facharzt verlängert sich die Befundungszeit durch CAD-Einsatz jedoch nicht signifikant. Interessanterweise benötigen die Befunder mit und ohne CAD-Anwendung für die einzelnen Patientengruppen unterschiedlich lange Befundungszeit. So befunden alle drei Befunder am schnellsten in der Früherkennungsgruppe. Der Assistenzarzt und der Student befunden am zweitschnellsten in der Benignengruppe, gefolgt von der Malignengruppe. Der Facharzt jedoch benötigt für die Befundung der Malignengruppe etwas weniger Zeit als für die Benignengruppe.

5.4 Verteilung der richtig-positiven und falsch-positiven Markierungen in Abhängigkeit vom jeweiligen Patientenkollektiv

In diesem Kapitel soll ermittelt werden ob es Seiten und Ebenenabhängige Markierungsunterschiede des CAD-Systems gibt.

Tabelle 10: Falsch-positive Markierungen in der Früherkennungsgruppe (n = 103): Ebenenabhängig, im Seitenvergleich und getrennt für Herdbefunde und Mikrokalk

Befund	Links		Rechts	
	CC	MLO	CC	MLO
Herdbefund pro Bild	0.3	0.4	0.4	0.4
Mikrokalk pro Bild	0.2	0.2	0.2	0.2
Σ pro Bild	0.5	0.6	0.6	0.6
Σ pro Seite	1.1		1.2	

Die Tabelle 10 stellt die Anzahl der falsch-positiven Markierungen der Früherkennungsgruppe dar. Diese werden Ebenen und Seitenabhängig in CC und MLO, sowie links und rechts dargestellt. Die Rate an falsch-positiven Markierungen ist im Ebenenvergleich mit 1,1 Markierungen in der CC-Ebene und 1,2 Markierungen in der MLO-Ebene fast gleich. Bei Betrachtung der Brustseitenverhältnisse kommen gleiche Ergebnisse zustande. Zusammenfassend markiert das CAD-System suspektere Areale Seiten und Ebenenunabhängig. Insgesamt wurden 2,3 falsch-positiv Markierungen gesetzt, wobei im Verhältnis mit 1,2 Markierungen mehr Herdbefunde als Mikroverkalkungen mit 1,1 Markierungen erfolgten.

Tabelle 11: Falsch-positive Markierungen in der Benignengruppe (n = 102): Ebenenabhängig im Seitenvergleich und getrennt für Herdbefunde und Mikrokalk

Befund	Links		Rechts	
	CC	MLO	CC	MLO
Ebene				
Herdbefund pro Bild	0.4	0.4	0.4	0.5
Mikrokalk pro Bild	0.6	0.4	0.5	0.5
Σ pro Bild	1.0	0.8	0.9	1.0
Σ pro Seite	1.8		1.9	

In der Tabelle 11 sind die falsch-positiven Markierungen für die Benignengruppe zusammengestellt. Hier werden erneut die falsch-positiven Markierungen ebenengetrennt und im Seitenvergleich betrachtet. Die Verteilung der falsch-positiven Markierungen im Hinblick auf die Projektionsebene, sowie im Seitenvergleich zeigt keinen signifikanten Unterschied. Auch hier markiert das CAD-System suspektere Areale Seiten und Ebenenunabhängig. Insgesamt wurden 3,7 falsch-positive Markierungen gesetzt. Innerhalb dieser falsch-positiven Markierungen waren mit 2 Markierungen mehr falsch-positive für Mikrokalk als mit 1,7 Markierungen für Herdbefunde gesetzt worden.

Tabelle 12: Falsch-positive Markierungen in der Malignengruppe (n = 98): Ebenenabhängig im Seitenvergleich und getrennt für Herdbefunde und Mikrokalk

Befund	Links		Rechts	
	CC	MLO	CC	MLO
Ebene				
Herdbefund pro Bild	0.4	0.3	0.2	0.4
Mikrokalk pro Bild	0.3	0.3	0.3	0.3
Σ pro Bild	0.7	0.6	0.5	0.7
Σ pro Seite	1.3		1.2	

Tabelle 13: Richtig-positive Markierungen in der Malignengruppe (n = 98): Ebenenabhängig im Seitenvergleich und getrennt für Herdbefunde und Mikrokalk

Befund	Links		Rechts	
	CC	MLO	CC	MLO
Herdbefund pro Bild	0.3	0.2	0.1	0.1
Mikrokalk pro Bild	0.2	0.2	0.3	0.3
Σ pro Bild	0.5	0.4	0.4	0.4
Σ pro Seite	0.9		0.8	

In Tabelle 12 bis Tabelle 13 ist die Verteilung der falsch-positiven, sowie der richtig-positiven Markierungen in der Malignengruppe aufgeführt. Hier erfolgt ebenfalls der Vergleich auf die prozentuale Verteilung der Markierungen bezogen auf die jeweilige Projektionsebene CC und MLO sowie beider Seiten links und rechts. Im Durchschnitt setzt das CAD-System in der CC-Ebene der linken Brust 0,7 Markierungen mehr falsch-positive Markierungen als in der MLO-Ebene mit 0,6 Markierungen. Auf der rechten Brustseite hingegen werden in der MLO-Ebene mit 0,7 Markierungen mehr falsch-positive Markierungen als in der CC-Ebene mit 0,5 Markierungen gesetzt. Die Anzahl der falsch-positiven Markierungen bezogen auf beide Seiten und Projektionsebenen sind mit 1,3 (links) zu 1,2 (rechts) Markierungen annähernd gleich. Die Verteilung der falsch-positiven Markierungen im Hinblick auf die Projektionsebene, sowie im Seitenvergleich zeigt keinen entscheidenden Unterschied. Bei Betrachtung der richtig-positiven Markierungen werden auf der rechten Brust mit 0,6 mehr Markierungen für Mikrokalk als links mit 0,4 durch das CAD-System gesetzt. Die Anzahl der richtig-positiven Markierungen für Herdbefunde ist auf der linken Seite mit 0,5 höher als auf der rechten Seite mit 0,2. Es werden im Verhältnis auf der linken Brustseite mit 0,9 Markierungen mehr richtig-positive Markierungen gesetzt als rechts mit 0,8 Markierungen. Insgesamt hat das CAD-System mit einer Trefferquote von 92% fast alle Tumore richtig-positiv erkannt.

6. Bilderteil

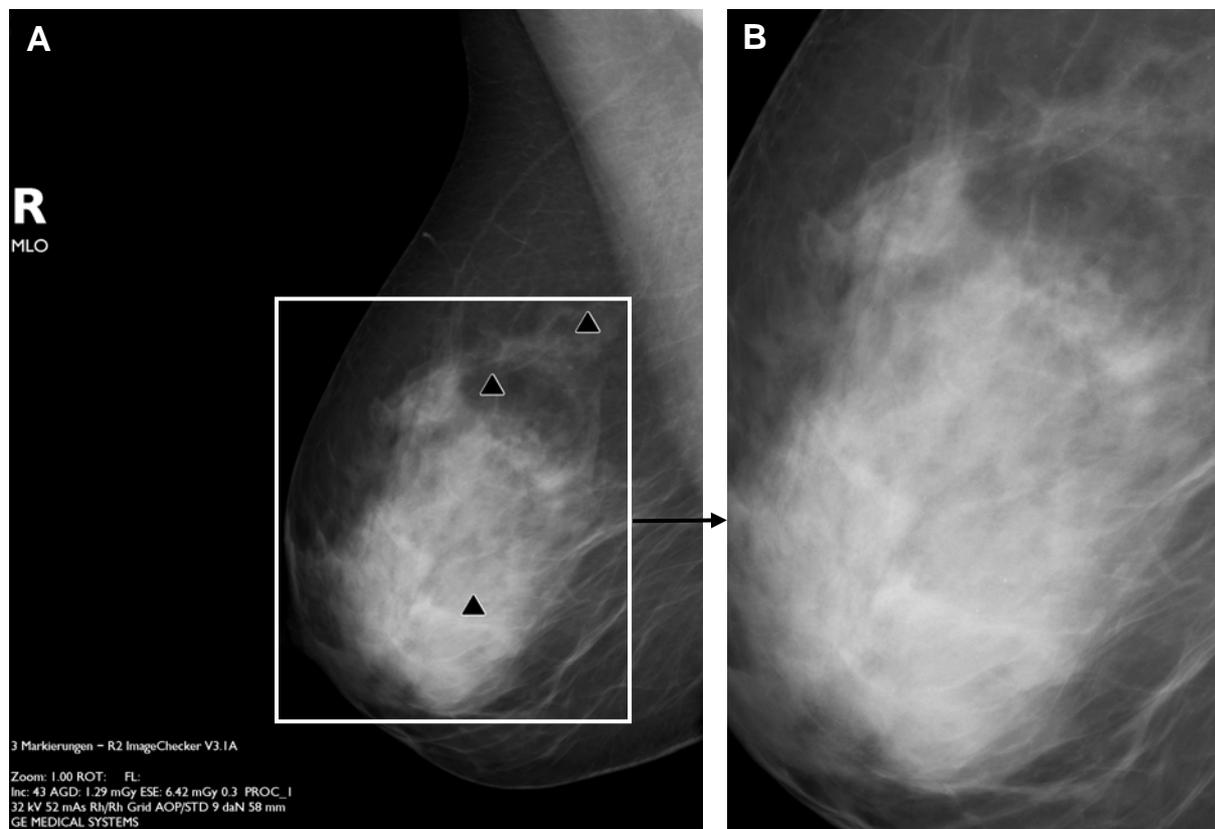


Abbildung 14: Darstellung der rechten Brust einer 66-jährigen Patientin aus der Benignengruppe. (A) In der MLO-Ebene drei falsch-positive Markierungen für Mirkokalk. (B) Vergrößerte Darstellung der drei falsch-positiven vaskulären Mirkokalzifikationen.

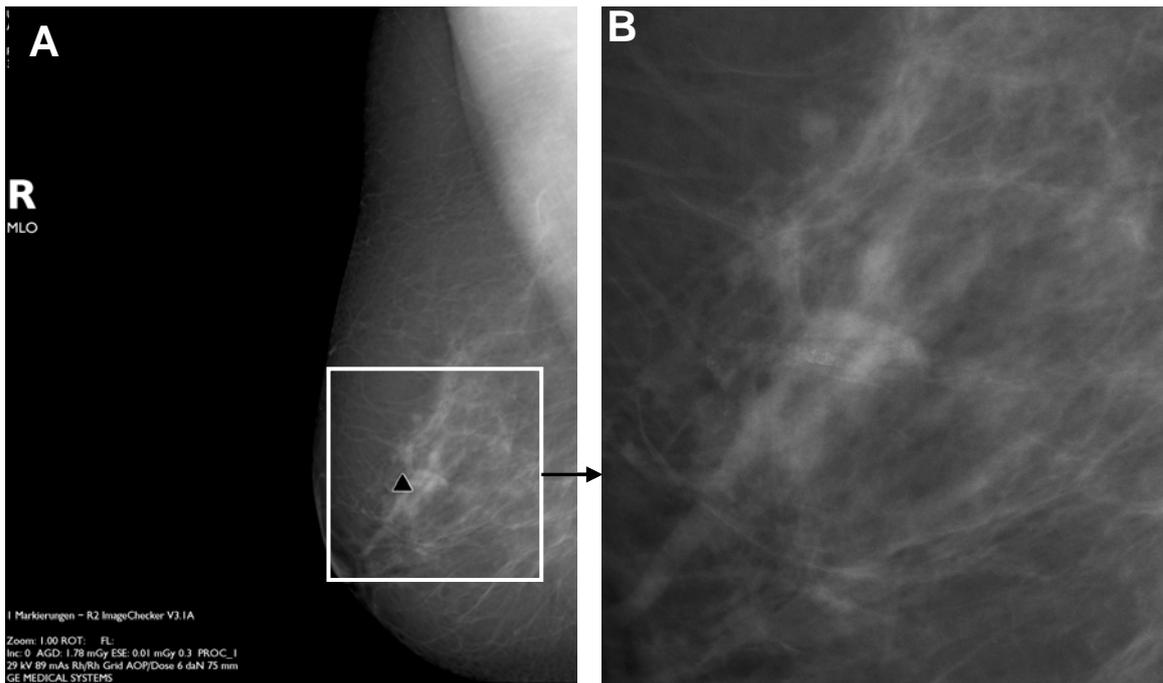


Abbildung 15: Darstellung der rechten Brust einer 51-jährigen Patientin aus der Früherkennungsgruppe. (A) In der MLO-Ebene eine falsch-positive Markierung für Mikrokalk. (B) Vergrößerte Darstellung.

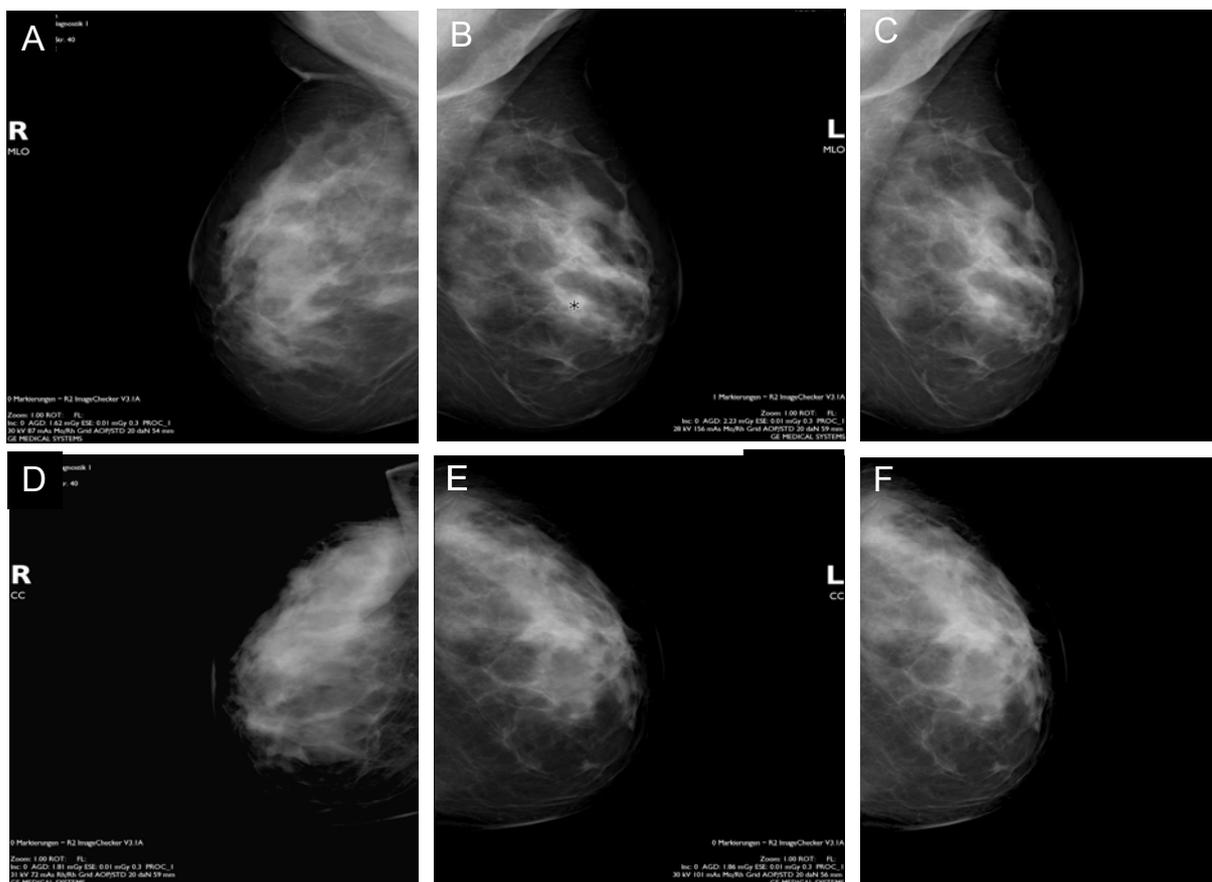


Abbildung 16: 65 Jahre alte Patientin aus der Früherkennungsgruppe. (A) In der MLO-Ebene der rechten Brust wurden keine Markierungen gesetzt hingegen eine falsch-positive Markierung für Herdbefund in der linken Brust (B). (C) Darstellung der entsprechenden linken MLO-Ebene ohne CAD-Markierungen. In der CC-Ebene der rechten Brust (D) und der linken Brust (E,F) wurden keine Markierungen gesetzt.

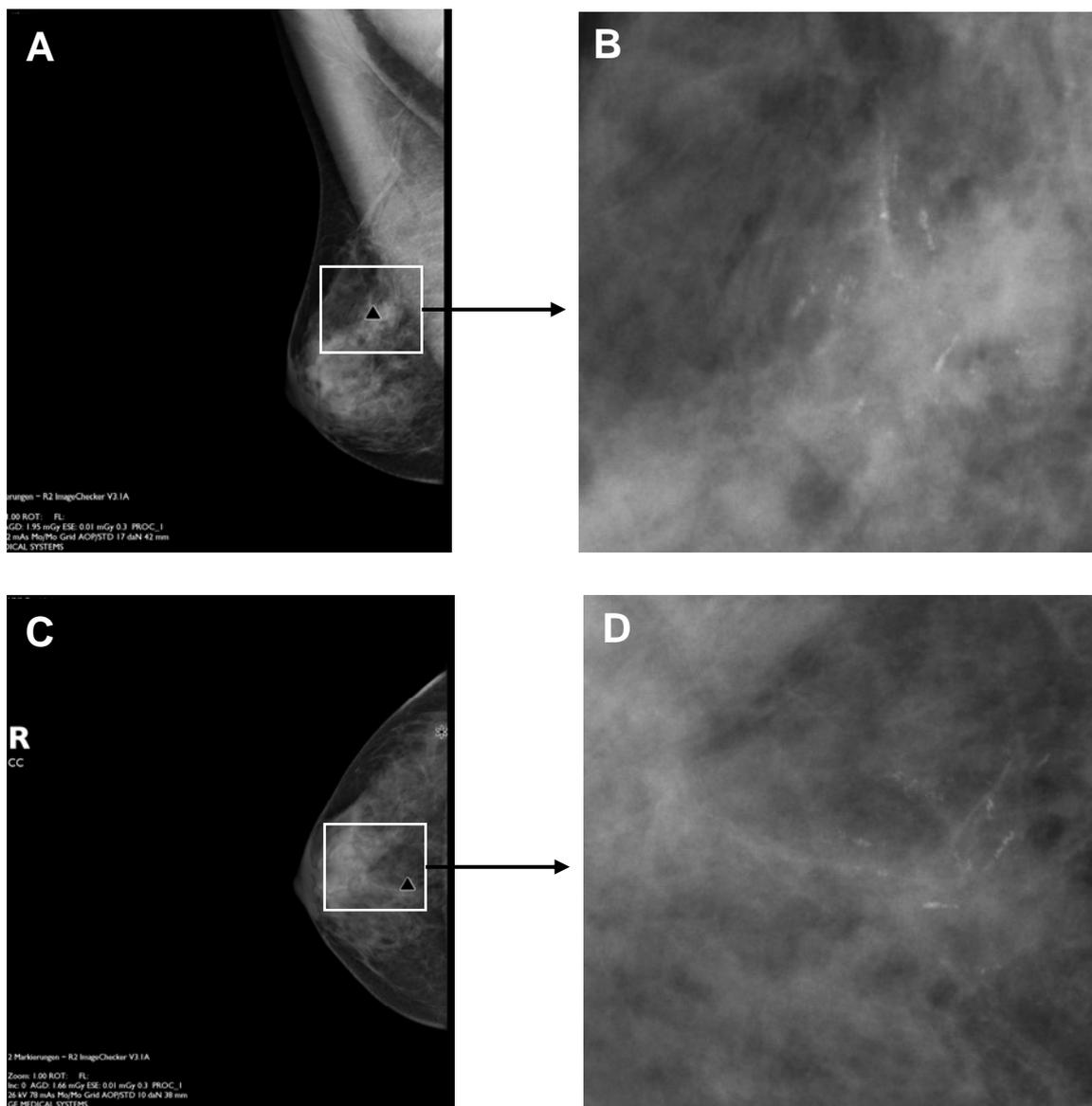


Abbildung 17: 55 Jahre alte Patientin mit histologisch gesichertem invasiv-duktalem Karzinom in der rechten Brust. (A) Richtig-positiv markierter Mikrokalk in der MLO-Ebene. (B) Vergrößerte Darstellung. (C) Richtig-positiv markierter Mikrokalk in der CC-Ebene, ebenfalls mit vergrößerter Darstellung (D). Zusätzlich eine falsch-positive Markierung eines Herdbefundes anhand der CC-Ebene.

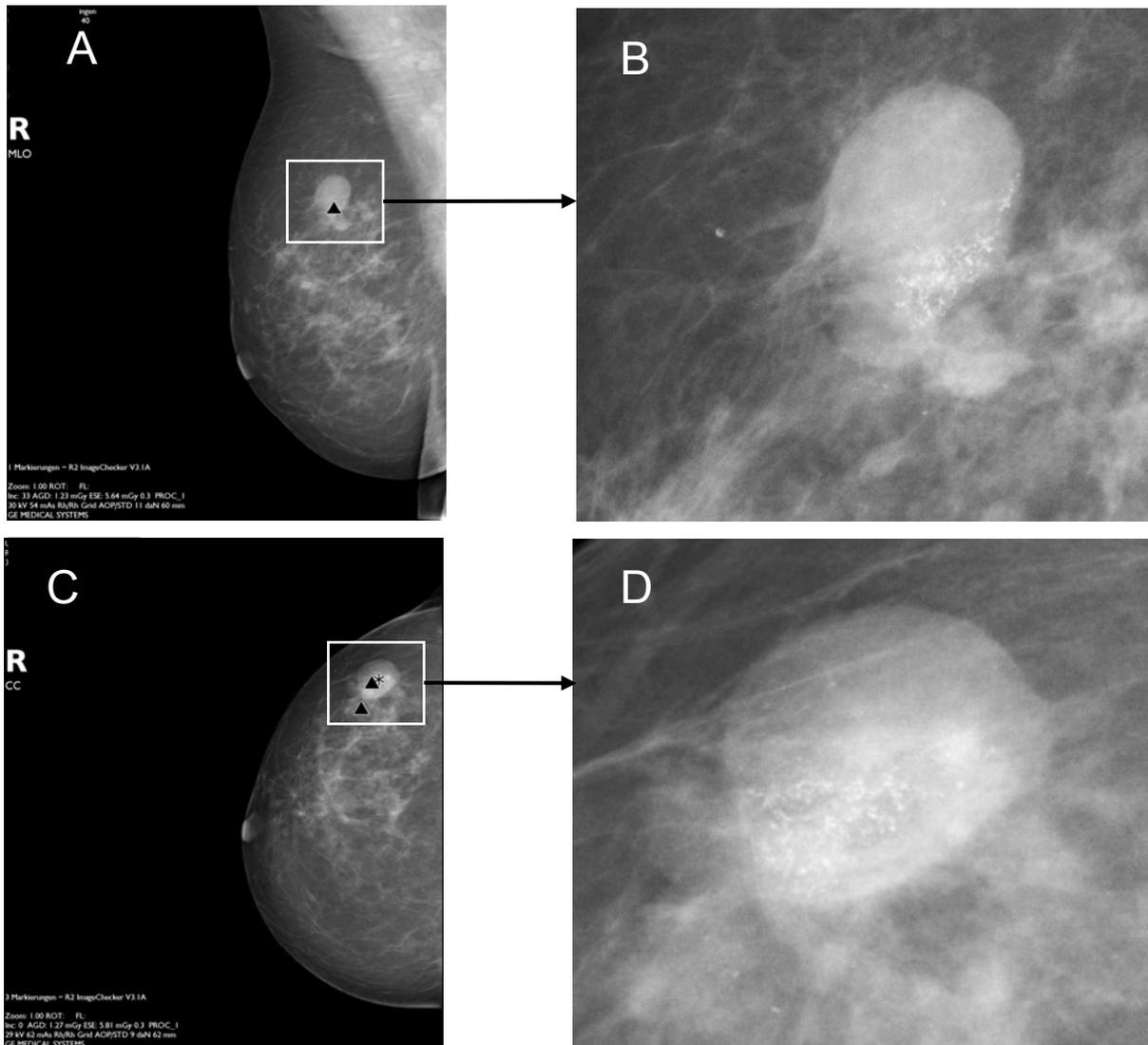


Abbildung 18: 55 Jahre alte Patientin mit histologisch gesichertem invasiv-duktalem Karzinom in der rechten Brust. (A) In der rechten MLO-Ebene ist eine richtig-positive Markierung für Mikrokalk gesetzt worden. (B) Vergrößerte Darstellung der Region mit Mikrokalk. (C) In der rechten CC-Ebene wurden zwei richtig-positive Markierungen für Mikrokalk gesetzt. In der CC-Ebene wurde eine richtig-positive Markierung für Herdbefunde gesetzt. (D) Vergrößerte Darstellung der Region mit Mikrokalk.

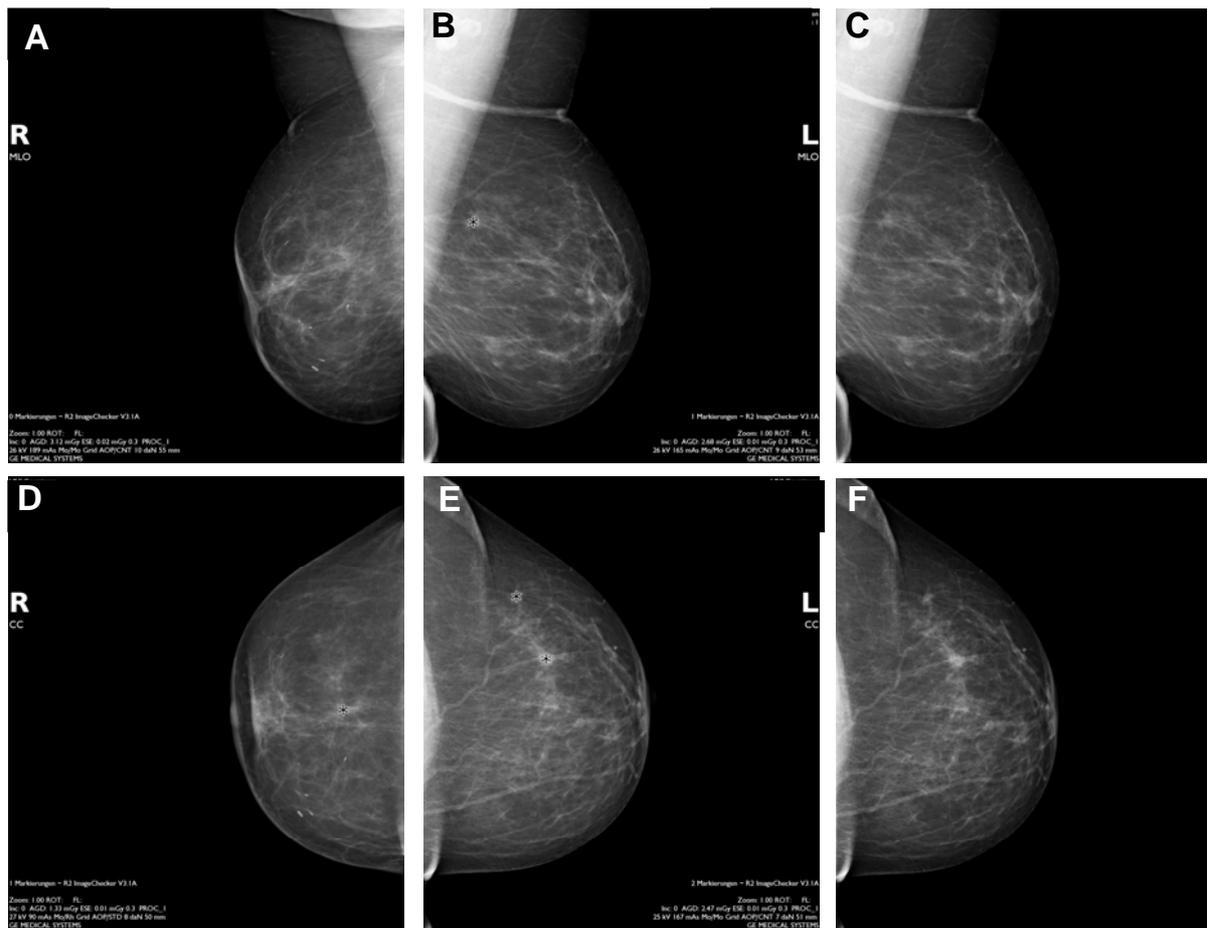


Abbildung 19: 68 Jahre alte Patientin mit positivem Tastbefund links außen und histologisch gesichertem invasiv-duktalem Karzinom. (A) In der rechten MLO-Ebene wurden keine Markierungen gesetzt. (B) In der linken MLO-Ebene wurde eine richtig- positive Markierung für Herdbefund gesetzt. (C) Im Vergleich die linke MLO-Ebene ohne CAD-Markierung. In der CC-Ebene eine falsch-positive Markierung in der rechten Brust (D) und in der linken Brust (E). Ebenfalls wurde eine richtig-positive Markierung für einen suspekten Herdbefund in der linken Brust gesetzt (E). Entsprechende CC-Ebene der linken Brust ohne CAD-Markierungen.

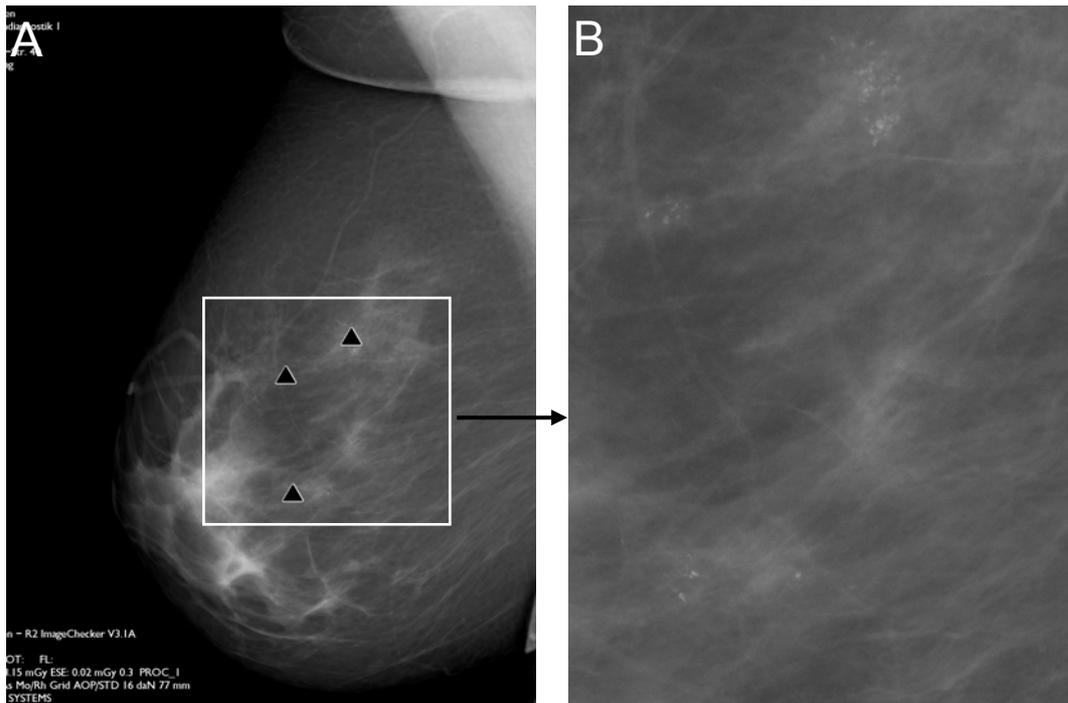


Abbildung 20: 66 Jahre alte Patientin aus der Früherkennungsgruppe. (A) In der MLO-Ebene drei falsch positive Markierungen für Mikrokalk. (B) Vergrößerte Darstellung der polymorphen Mikrokalzifikationen (BIRADS 4). Histologie: Duktale Hyperplasie mit Mikrokalzifikationen.

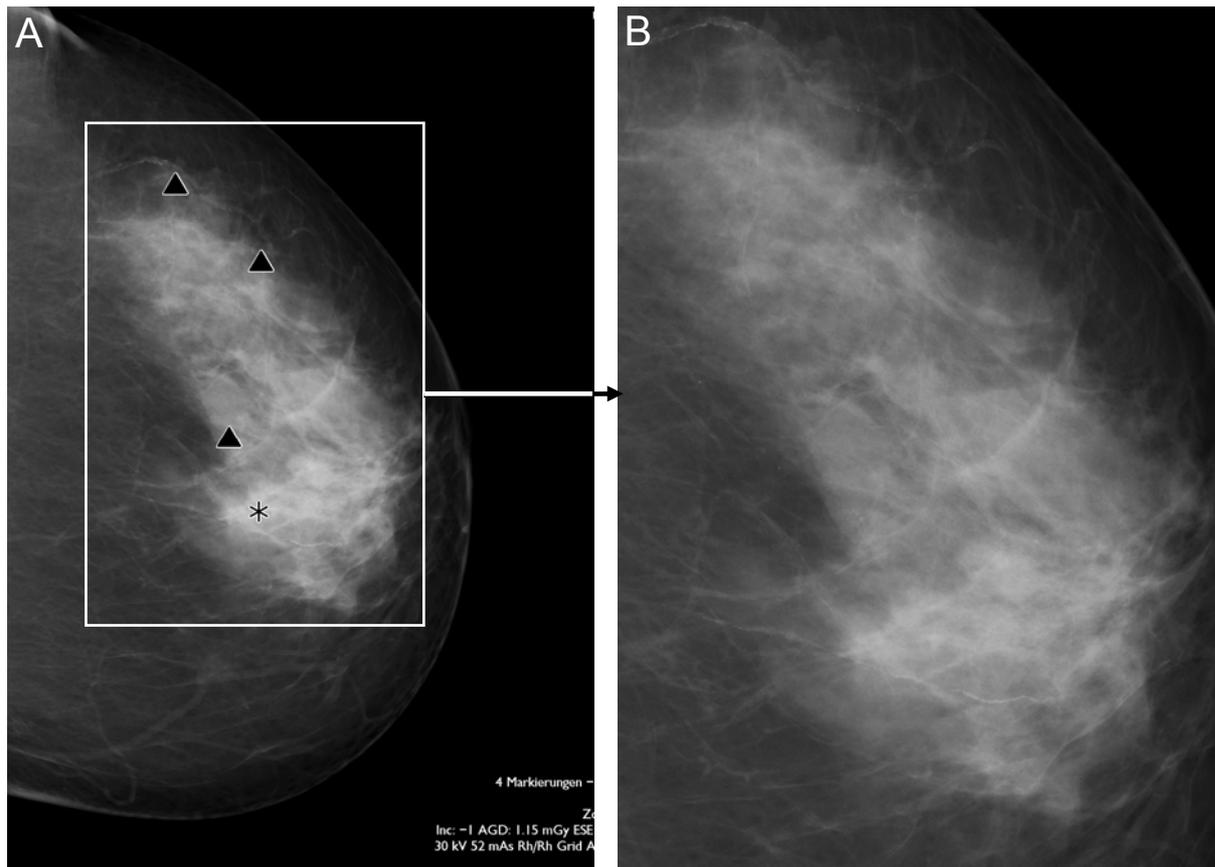


Abbildung 21: Darstellung der linken Brust einer 65-jährigen Patientin aus der Benignengruppe. (A) In der linken CC-Ebene drei falsch positive Markierungen für Mikrokalk und eine falsch positive Markierung für Herdbefund. (B) Vergrößerte Darstellung.

7. Diskussion

Die Mammographie ist die Basis der bildgebenden Diagnostik des Mammakarzinoms und eine Methode, die nachgewiesenermaßen die Letalität dieser Erkrankung senkt (Shapiro et al. 1971; Tabar et al. 1985). Im klinischen Alltag wird die Mammographie als einziges effektives Verfahren im Rahmen der Brustkrebsfrüherkennung eingesetzt, was zu einer erhöhten Heilungswahrscheinlichkeit führt. Eine Verbesserung in der Detektion von Mammakarzinomen durch die Unterstützung von CAD-Systemen haben mehrere Studien bewiesen. Die Sensitivität konnte mithilfe der computerassistierten Diagnose um 12 bis 19,5% erhöht werden (Freer, Ulissey 2001; Quek et al. 2003; Cupples 2001; Bandodkar et al. 2002). Zur Bestätigung der oben genannten Ergebnisse gibt es weitere Studien, die an einem begrenzten Kollektiv die Genauigkeit von Radiologen mit und ohne CAD-Unterstützung an denselben Aufnahmen vergleichen (Jiang 1996; Chan et al. 1990). Es wird bereits heute diskutiert, dass CAD-Systeme zukünftig die Rolle des zweiten Radiologen übernehmen könnten (Obenauer et al. 2005; Thurfjell et al. 1994; Obenauer et al. 2006). Ein Problem beim Einsatz von CAD-Systemen stellt jedoch nach wie vor die häufige falsch-positive Markierung dar (Obenauer et al. 2005; Obenauer et al. 2006; Castellino, Roehring 1999, Baum et al. 2002; Malich et al. 2003). Dies trägt in entscheidendem Maße dazu bei, dass mit der Interpretation eines CAD-Befundes heute oftmals ein hoher Zeitaufwand verbunden ist, da der Radiologe von der Tumormarkierung des CAD-Systems abgelenkt werden kann. In unserer Studie ist eine CAD-abhängige Verlängerung der Befundungszeit signifikant nachgewiesen (Tabelle 6 bis Tabelle 9; Abbildung 13). Der Facharzt mit langjähriger Mammographieerfahrung benötigt erwartungsgemäß, mit und ohne Unterstützung des CAD-Systems, die kürzeste Zeit zur Befundauswertung. Auffällig ist, dass der erfahrene Befunder im Median mit und ohne CAD-Unterstützung die annähernd gleiche Befundungszeit benötigt. Auch ist die zeitliche Streuung kaum unterschiedlich. Er scheint am wenigsten vom CAD-System zu profitieren. Der Assistenzarzt benötigt deutlich mehr Zeit für die Befundung. Auch die Spannweite von maximaler und minimaler Zeit ist deutlich größer als beim Facharzt. Die längste Zeit für die Mammographiebefundung verwendet der Student. Der CAD-Einsatz verlängert die Zeit zusätzlich. Auch die mediane Befundungszeit hat sich durch CAD-Anwendung verlängert. Wie zuvor beschrieben, ist der entscheidende Faktor für die Verlängerung der Befundungszeit die häufigen falsch-positiven Markierungen. Dies muss jedoch mit dem positiven Effekt der Sensitivitätszunahme abgewogen werden.

Das Ausmaß der zu erwartenden CAD-Effekte wird vom Sensitivitätsniveau und der Rate an falsch-positiven Befunden abhängen. Da bei der Screeningmammographie die Sensitivität bei

ca. 85% liegt (Baines et al. 1986, Bird et al. 1992), ist ein relativ großer Spielraum für CAD-assoziierte Verbesserungen gegeben. Bei höheren Sensitivitäten von erfahrenen Radiologen hingegen, sind nur wenige zusätzliche richtig-positive Markierungen zu erwarten, wenn ein solcher Radiologe das CAD-System zusätzlich einsetzt. Daraus ergibt sich, dass der maximale CAD-Effekt (bei Berücksichtigung aller richtig-positiven Markierungen) umso größer ist, je geringer die Erfahrungen des Radiologen sind. Dementsprechend müsste der Radiologe mit der niedrigsten Sensitivität am meisten von der Computerhilfe profitieren. Daher bietet es sich aus diesem Grund zum Beispiel an, CAD-Systeme als Lernhilfe in der Ausbildung einzusetzen. Der lernende Radiologe erkannte in einer Studie von Quek et al. (2003) von 294 Tumoren nur 257 (87,4%) selbstständig, bei 21 Tumoren (7,2%) assistierte das CAD-System bei der Detektion, in 13 Fällen (4,4%) deutete das CAD auf einen möglichen Befund hin, der von diesem Radiologen nicht erkannt wurde, und in 3 Fällen (1%) sorgte das CAD mit einer falsch-positiven Markierung für Irritationen. Insgesamt wurde in der Studie die Sensitivität der unerfahrenen Radiologen durch den Einsatz des CAD-Systems von 74,4% (ohne CAD) auf 87,2% signifikant gesteigert (Quek et al. 2003). Ein weiteres Kriterium für die Befundungszeit stellt das Patientenkollektiv dar. In der Früherkennungsgruppe, in der nur Patientinnen mit sicher benignen Befunden zusammengefasst wurden, gab es keine signifikant unterschiedlichen Befundungszeiten ($p= 0.1343$, Tabelle 7, Abbildung 13). Bezüglich der Befundererfahrung gibt es allerdings einen deutlich signifikanten Unterschied des Zeitaufwands ($p < 0.0001$, Tabelle 7, Abbildung 13). Der Einfluss auf die Zeit bei Kombination von Befundererfahrung und dem Einsatz der CAD-Technik ist nicht signifikant unterschiedlich (Tabelle 7). Im Vergleich dazu liefert der p-Wert bei der Malignengruppe (Tabelle 9) und bei der Benignengruppe (Tabelle 8) durchweg signifikante Ergebnisse bezüglich der Befundungszeit mit und ohne CAD-Einsatz, der Befundererfahrung und der Kombination von Befundererfahrung und CAD-Unterstützung. Der zukünftige CAD-Einsatz im klinischen Alltag wird in erster Linie davon abhängen, ob die Sensitivität bei der Erkennung von malignen Befunden in der Mammographie weiter gesteigert werden kann und die Rate an falsch-positiven Markierungen durch Weiterentwicklungen der CAD-Systeme und der dazugehörigen Software vermindert werden kann. Neue CAD-Systeme auf der Basis neuronaler Netze haben die zusätzliche Fähigkeit der automatisierten permanenten Datenakquirierung als Datenpool zum Abgleich neuer Mammographien. Diese Systeme werden zurzeit wissenschaftlich untersucht (Elter et al. 2007). Abschließend kommt man zu dem Ergebnis, dass trotz der oben genannten Einschränkungen das CAD-System bereits heute eine wesentliche Hilfe bei der Interpretation von Mammographien darstellt und den Zweitbefunder

bei der Mammographiebefundung ersetzen könnte. Dies ist jedoch entscheidend davon abhängig wie umfangreich die Erfahrung des Erstbefunders ist. Die meisten klinischen Studien, in denen mögliche Nutzen der computer-assistierten Diagnose (CAD) als Unterstützung für Radiologen in ihren Mammographiebefundungen erfolgte, waren retrospektiv. Viele Publikationen haben das Potenzial von CAD gezeigt, indem sie die Anzahl falsch-negativer Befundungen verringerten, indem sie die CAD-Analyse retrospektiv auf Mammographien bereits erkrankter Brustkrebspatienten durchführten. Aufgrund der Tatsache, dass es dem CAD-System gelingt, retrospektiv maligne Strukturen rückblickend zu markieren, schätzten einige Studien den möglichen Nutzen des CAD-Systems, auf eine Zunahme der Brustkrebsdetektionsrate von 21% ein (Brem et al. 2003). Das Ziel unserer Studie war es, festzustellen, welche Auswirkungen ein kommerziell verfügbares CAD-Systems auf die Interpretationen der drei unterschiedlich erfahrenen Befunder bezogen auf die unterschiedlichen 3 Patientenpopulationen hat. Die Studiengruppen Freer und Ulissey haben 12 800 Mammographie retrospektiv mit dem ImageChecker erfasst. Die Radiologen erkannte 41 Karzinome ohne CAD-Unterstützung. Unter Zuhilfenahme des CAD-Systems wurden acht zusätzliche Karzinome erkannt, von denen sieben Mikrokalzifikationen aufwiesen. Diese Regionen wurden erst durch den Radiologen überprüft, nachdem diese Areale vom CAD-System markiert wurden. Der Einsatz des CAD-Systems führte in dieser Studie zu einem Anstieg in der Brustkrebs-Erkennungsrate von 19,5% (Freer, Ulissey 2001). Dieses Ergebnis ist vergleichbar mit anderen retrospektiven Studien, die auch einen Anstieg der Brustkrebs-Erkennung von ca. 20% erzielten (Burhenne HJ et al. 2000). In der Studie von Freer und Ulissey (2001) entdeckten die Reader 84% der gesamten Mammakarzinome ohne Zuhilfenahme des CAD-Systems. Das CAD-System markierte 82% der Mammakarzinome richtig-positiv. Die Forschungsgruppe Morton et al. (2006) veröffentlichte eine weitere große prospektive Studie in der die Effektivität des CAD-System im klinischen Alltag untersucht wurde. In dieser Studie wurden 21 349 Patientinnen einer Mammographie unterzogen. Die Radiologen diagnostizierten 105 Mammakarzinome ohne die Unterstützung des CAD-Systems. Unter Zuhilfenahme des CAD-Systems wurden 8 zusätzliche Mammakarzinome entdeckt, die vorher von den Radiologen übersehen worden waren. Insgesamt wurden 113 Mammakarzinome mit der Unterstützung des CAD-Systems diagnostiziert, was einem Anstieg der Detektionsrate von 7.62% entspricht. Dieses Ergebnis unterscheidet sich von den Ergebnissen anderer retrospektiven Untersuchungen, dort kam es zu einer Zunahme von ca. 20% in der Detektionsrate, unter Einsatz des CAD-Systems (Burhenne HJ et al.2000).

In unserer Studie konnten wir zeigen, dass alle drei Befunder eine Steigerung ihrer Treffgenauigkeit mit Hilfe des CAD-Systems erreichten, unabhängig der untersuchten 201 Patientinnen die sich aus der Malignengruppe (n=98) und Früherkennungsgruppe (n=103) zusammensetzten. Alle drei Befunder konnten eine Steigerung der AUC erreichen, wenn sie das CAD-System einsetzten (Abbildung 10 und Abbildung 12, Tabelle 2 und Tabelle 4). Tabelle 2 zeigt, dass alle drei Befunder eine höhere AUC erreichten, wenn sie für die Befundung der Malignengruppe und der Früherkennungsgruppe das CAD-System einsetzten. Den größten Nutzen hat der Student unter Zuhilfenahme des CAD-Systems. Seine AUC steigt um 10% gefolgt vom Assistenzarzt mit 3% und schließlich vom Facharzt mit 2%. Im Durchschnitt steigt hier die AUC um 5%. Tabelle 4 zeigt isoliert die Auswertung der AUC der Malignengruppe, auch hier steigt die AUC mit dem CAD-System. Die höchste Steigerung der AUC erreichte auch hier der Student (10%), gefolgt vom Assistenzarzt (4%) und dem Facharzt (3%). Unsere Studie zeigt signifikante Unterschiede der AUC im Hinblick auf die Untersuchererfahrung und die Methode mit/ohne CAD (Tabelle 2 und Tabelle 4). Es scheint, dass die jeweilige Patientengruppe keinen signifikanten Einfluss auf die Zunahme der AUC besitzt, wenn das CAD-System verwendet wird (Tabelle 3). Alle drei Befunder haben fast dieselbe Zunahme der AUC in den Befundungen der Malignengruppe und der Früherkennungsgruppe, wenn sie das CAD-System anwandten. Ein weiterer Aspekt der Arbeit war es, welche Rolle das CAD-System einnimmt, wenn Befunder mit unterschiedlicher Erfahrung dieses System einsetzen. Eine Studie durch Quek et al. (2003) evaluierte die Rolle des CAD-Systems für die Entdeckung von Mammakarzinomen durch unerfahrene Reader. Die Mammographien wurden von unerfahrenen Readern zuerst ohne und dann mit Hilfe des CAD-Systems bewertet. Später wurden die Mammographien durch erfahrene Reader bewertet, diese bildeten die Referenzgruppe. Insgesamt führte das CAD-System zu einer signifikanten Steigerung der Sensitivität von 74,4% auf 87,2%. Das CAD-System verbesserte die Detektionsrate unerfahrener Befunder, diese waren dadurch in der Lage suspekta Areale besser zu identifizieren. Die Tatsache, dass die Sensitivität unerfahrener Befunder in Kombination mit dem CAD-System steigt, zeigt, dass in Zukunft ein Einsatz solcher Systeme in der radiologischen Ausbildung möglich ist. In unserer Studie hatte der unerfahrene Student den größten Anstieg seiner AUC, wenn er das CAD-System einsetzte. Er steigerte im Schnitt seine AUC um 10% (Tabelle 2, Tabelle 4). Der Einfluss der Berufserfahrung der Befunder ist in unserer Studie signifikant (Tabelle 3, Tabelle 5). Jedoch ist es dem ungeübten Student nicht möglich eine AUC wie der Assistenzarzt oder Facharzt zu erreichen, auch wenn er das CAD-System einsetzt und dem Assistenzarzt ist es nicht möglich

Facharzniveau mit Einsatz des CAD-Systems zu erreichen (Abbildung 10 und 12). Das CAD-System ist als ein potentieller Ersatz für das heutige *double-reading* vorgeschlagen worden (Ciatto et al. 2004). Das *double-reading* im Mammographiescreening erhöht die Tumordetektionsrate, so dass zwischen 3-15 Tumoren pro 100000 Frauen früher entdeckt und therapiert werden könnten, abhängig davon welche Methode des *double-readings* durchgeführt wird (Dinnes et al. 2001). In Europa ist das *double-reading* aufgrund der Sensitivitätssteigerung Standard, nicht jedoch in den USA, größtenteils aufgrund der hohen finanziellen und logistischen Belastung (Elmore, Carney 2004). Heute können CAD-Systeme in die tägliche Routine integriert werden, aber ihre zukünftige Rolle hängt von ihrer weiteren Entwicklung ab. Schließlich zeigt unsere Studie, dass die Zunahme der AUC vom Erfahrungsstand der Befunder abhängt. In allen untersuchten Gruppen verursachte die Anwendung des CAD-Systems eine steile Zunahme der ROC-Kurve und eine steilere AUC und folglich eine Steigerung der Sensitivität. Die Reduzierung der falsch-positiven Markierungen des CAD-Systems ist die grundlegende Voraussetzung, um das *double-reading* eines Tages in der klinischen Arbeit zu ersetzen.

8. Zusammenfassung

Das Ziel dieser Arbeit war es, zu evaluieren, welchen Einfluss die Anwendung des computerassistierten Diagnosesystems, der Ausbildungsstand der Befunder und das jeweilige Patientenkollektiv auf die Befundungszeit, Trefferquote und Anzahl der falsch-positiven sowie falsch-negativen Markierungen von Mammographieaufnahmen haben. Das Patientenkollektiv bestand aus 303 Patientinnen, die in drei Evaluationsgruppen unterteilt wurden: 1) Früherkennungsgruppe: 103 Patientinnen, 2) Benignengruppe: 102 Patientinnen, 3) Malignengruppe: 98 Patientinnen. Ein Assistenzarzt mit fortgeschrittenem Ausbildungsstand, ein Student und ein Facharzt mit Spezialkenntnissen in der Mammographie haben die einzelnen Mammographieaufnahmen in zufälliger Reihenfolge und ohne die Kenntnisse der jeweiligen Patientengruppe oder Vorbefunde, mit und ohne CAD-Unterstützung ausgewertet. Die Ergebnisse der Befunder wurden vermerkt, um statistisch die Sensitivität und Spezifität zu ermitteln. Um die Sensitivität und Spezifität der drei Befunder mit und ohne CAD-System zu bewerten, wurde unter Berücksichtigung der Befundungszeiten, Untersuchererfahrung und der Tumormanifestation für jeden einzelnen Befunder eine ROC-Analyse mit der entsprechenden Fläche unter der Kurve AUC (AUC = area under the curve) erstellt. Um das CAD-System zu überprüfen, wurden alle gesetzten Markierungen notiert und mit den Mammographien verglichen. Die Markierungen wurden in falsch-positiv und richtig-positiv unterteilt. In der Früherkennungsgruppe (siehe Tabelle 11) war die Rate an falsch-positiven Markierungen im Ebenenvergleich mit 1,1 Markierungen in der CC-Ebene und 1,2 Markierungen in der MLO-Ebene fast gleich. Bei Betrachtung der richtig positiven Markierungen werden auf der rechten Brust mit 0,6 mehr Markierungen für Mikrokalk als links mit 0,4 durch das CAD-System gesetzt. Die Anzahl der richtig-positiven Markierungen für Herdbefunde ist auf der linken Seite mit 0,5 höher als auf der rechten Seite mit 0,2. Es werden im Verhältnis auf der linken Brustseite mit 0,9 Markierungen mehr richtig-positive Markierungen gesetzt als rechts mit 0,8 Markierungen. Insgesamt hat das CAD-System mit einer Trefferquote von 92 % fast alle Tumore richtig-positiv erkannt. Die ROC-Kurven zeigten, dass der Student mit und ohne CAD die flachste Zunahme der Kurve und infolgedessen die geringste Sensitivität und Spezifität besitzt. Mit der Unterstützung des CAD-Systems ist die ROC-Kurve steiler. Der Assistenzarzt erreicht mit und ohne den Einfluss von CAD einen steileren Anstieg des Diagramms als der Student. Unter der Zuhilfenahme des CAD-Systems ist die Kurve auch steiler als ohne CAD-Anwendung. Dies zeigt, dass der Assistenzarzt eine höhere Sensitivität und Spezifität als der Student erzielt. Der steilste Anstieg der ROC-Kurve und infolgedessen

die höchste Sensitivität und Spezifität hat der Facharzt mit und ohne den CAD-Einsatz. In dieser Studie haben alle drei Befunder einen Anstieg in ihrer Trefferquote, wenn sie das CAD-System verwenden. Den höchsten Nutzen hat der Student (10% Zunahme), gefolgt vom Assistenzarzt (4%) und schließlich gefolgt vom Facharzt (3%). Der Facharzt benötigt für die Befundung unabhängig vom CAD-Einsatz erwartungsgemäß die geringste Zeit, gefolgt vom Assistenzarzt und vom Studenten. In allen drei Patientenkollektiven verlängert die CAD-Anwendung die Befundungszeit des Studenten und des Assistenzarztes. Der Facharzt benötigt mit und ohne CAD-Technik im Median die gleiche Zeit. Bei der Früherkennungsgruppe zeigten sich keine signifikant unterschiedlichen Befundungszeiten ($p=0.1343$). Bezüglich der Befundererfahrung gibt es einen signifikanten Unterschied ($p<0.0001$). Der Einsatz der CAD-Technik korreliert mit der jeweiligen Befundererfahrung und liefert ebenfalls ein nicht signifikantes Ergebnis. Im Vergleich dazu liefert der p-Wert bei der Malignen- und bei der Benignengruppe signifikante Wechselwirkung zwischen den Erfahrungswerten der drei Befunder sowie dem Einsatz der CAD-Technik. Die Reduzierung der falsch-positiven Markierungen des CAD-Systems ist die grundlegende Voraussetzung, um das *double-reading* eines Tages in der klinischen Arbeit zu ersetzen.

Anhang

Literaturverzeichnis

Baines CJ, Miller AB, Wall C, McFarlane DV, Simor IS, Jong R, Shapiro BJ, Audet L, Petitiere M, Ouimet-Oliva D et al. (1986): Sensitivity and specificity of first screen mammography in the Canadian national breast screen study: A preliminary report from five centers. *Radiology* 160, 295-298

Bandodkar P, Birdwell RL, Ikeda DM (2002): Computer aided detection (CAD) with screening mammography in an academic institution: preliminary findings. *Radiology* 225, 458

Baum F, Fischer U, Obenaus S, Grabbe E (2002): Computer-aided detection in direct digital full-field mammography: initial results. *Eur Radiol* 12, 3015-3017

Bird RE, Wallace TW, Yankaskas BC (1992): Analysis of cancers missed at screening mammography. *Radiology* 3, 613-617

Brem RF, Baum J, Lechner M, Kaplan S, Souders S, Naul LG, Hoffmeister J (2003): Improvement in sensitivity of screening mammography with computer-aided detection: a multiinstitutional trial. *AJR Am J Roentgenol* 181, 687-693

Burhenne HJ, Burhenne LJ, Goldberg F, Hislop TG, Worth AJ, Rebbeck PM, Kan L (1994): Interval breast cancers in the Screening Mammography Program of British Columbia: analysis and classification. *AJR Am J Roentgenol* 5, 1067-1071

Burhenne LJ, Wood SA, D'Orsi CJ, Feig SA, Kopans DB, O'Shaughnessy KF, Sickles EA, Tabar L, Vyborny CJ, Castellino RA (2000): Potential contribution of computer-aided detection to the sensitivity of screening mammography. *Radiology* 215, 554-562

Castellino RA, Roehring JR (1999): The promise of computer aided detection in digital mammography. *Eur Radiol* 31, 35-39

Chan HP, Doi K, Vyborny CJ, Schmidt RA, Metz CE, Lam KL, Ogura T, Wu YZ, MacMahon H (1990): Improvement in radiologists detection of clustered microcalcifications on mammograms. The potential of computer-aided diagnosis. *Invest Radiol* 10, 1102-1110

Ciatto S, Ambrogetti D, Bonardi R, Brancato B, Catarzi S, Risso G, Rosselli Del Turco M (2004): Comparison of two commercial systems for computer-assisted detection (CAD) as an aid to interpreting screening mammograms. *Radiol Med* 107, 480-488

Cupples TE (2001): Impact of computer-aided detection (CAD) in a regional screening mammography program. *Radiology* 221, 221-520

Diekmann S, Diekmann F (2008): Mammography screening in Germany. *Radiologe* 1, 17-25

Dinnes J, Moss S, Melia J, Blanks R, Song F, Kleijnen J (2001): Effectiveness and cost-effectiveness of double reading of mammograms in breast cancer screening: findings of a systematic review. *Breast* 6, 455-463

D'Orsi CJ, Basset LW, Berg et al.: American College of Radiology, Reston, VA, (1998) *Breast Imaging Reporting and Data System: ACR BI-RADS-Mammography (ed3)*

Elmore JG, Carney PA (2004) Computer-aided detection of breast cancer: has promise outstripped performance? *J Natl Cancer Inst* 3, 162-163

Elter M, Schukz-Wendtland R, Wittenberg T (2007): The prediction of breast cancer biopsy outcomes using two CAD approaches that both emphasize an intelligible decision process. *Med Phys* 23, 4164-4172

Freer TW, Ulissey MJ (2001): Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology* 3, 781-786

Funovics M, Schamp S, Helbich TH, Lackner B, Wunderbaldinger P, Fuchsjäger M, Lechner G, Wolf G (2001): Evaluation of a computer-assisted diagnosis system in breast carcinoma. *Rofo* 3, 218-223

Giger ML (1993): Computer-aided Diagnosis. Technical aspects in Breast Imaging. *RSNA Oak Brook* 2, 283-298

Jiang Y (1996): Malignant and benign clustered microcalcifications: Automated feature analysis and classification. *Radiology* 198, 671-678

Konietschke F, Brunner E. (2009): Nonparametric analysis of clustered data in diagnostic trials: Estimation problems in small sample sizes. *Computational Statistics and Data Analysis* 53, 730 - 741.

Malich A, Sauner D, Marx C, Facius M, Boehm T, Pfliederer SO, Fleck M, Kaiser WA (2003): Influence of breast lesion size and histologic findings on tumor detection. Rate of a computer-aided detection system. *Radiology* 228, 851-856

Morton MJ, Whaley DH, Brandt KR, Amrami KK (2006): Screening mammograms: interpretation with computer-aided detection--prospective evaluation. *Radiology* 239, 375-338

Obenauer S, Hermann KP, Grabbe E (2003): Dose reduction in full-field digital mammography: an anthropomorphic breast phantom study. *Br J Radiol* 907, 478-482

Obenauer S, Sohns C, Werner C, Grabbe E (2005): Retrospective analysis of a computer-aided detection (CAD) system in full-field digital mammography in correlation to tumor histology. *Rofo* 177, 1103-1109

Obenauer S, Sohns C, Werner C, Grabbe E (2006): Computer-aided detection in full-field digital mammography: detection in dependence of the BI-RADS categories. *Breast J* 1, 16-19

Quek ST, Thng CH, Khoo JB, Koh WL (2003): Radiologists detection of mammographic abnormalities with and without a computer-aided detection system. *Australas Radiol* 3, 257-260

Shapiro S, Strax P, Venet L (1971): Periodic breast cancer screening in reducing mortality from breast cancer. *JAMA* 215, 1777-1785

Tabar L, Fagerberg CJ, Gad A, Baldetorp I, Holmberg LH, Grontoft O, Ljungquist U, Lundstrom B, Manson JC, Eklund G: Reduction in mortality from breast cancer after mass screening with mammography. *Lancet* 1985, 2, 829-832

Thurfjell EL, Lernevall KA, Taube AA (1994): Benefit of independent double reading in a population-based mammography screening program *Radiology* 1, 241-244

Werner, C. (2006). Nichtparametrische Analyse von diagnostischen Tests. Dissertation, Universität Göttingen

Werner, C, Brunner E, (2007): Rank methods for the analysis of clustered data in diagnostic trials. *Computational Statistics & Data Analysis* 51, 5041 - 054

Werner J: Urban & Schwarzenberg, München 1992: Biomathematik und Medizinische Statistik 2., überarbeitete Auflage

Tabellenverzeichnis

Tabelle 1: BI-RADS Klassifikation (D' Orsi et al. 1998): BI-RADS Klassifikation (D' Orsi et al. 1998).....	9
Tabelle 2: AUC aller drei Befunder in Abhängigkeit der Fläche unter der Kurve (AUC= area under the curve) für die Malignen - und Früherkennungsgruppe mit und ohne Anwendung des CAD-Systems.....	16
Tabelle 3: Unterschiede in der AUC für die Malignengruppe und Früherkennungsgruppe bezogen auf die Befunder, Methode und die Interaktion Befunder und Methode.....	16
Tabelle 4: AUC aller drei Befunder in Abhängigkeit der Fläche unter der Kurve für die Malignengruppe mit und ohne Anwendung des CAD-Systems:	18
Tabelle 5: Die unterschiedliche Treffgenauigkeit für die Malignengruppe und die Früherkennungsgruppe bezogen auf Befunder und Methode und Interaktion Befunder und Methode:.....	19
Tabelle 6: Abhängigkeit der Befundungszeiten vom jeweiligen Ausbildungsstand der Befunder:.....	20
Tabelle 7: Unterschiede in den Befundungszeiten in der Früherkennungsgruppe in Abhängigkeit vom Befunder sowie von der Anwendung des CAD-Systems:.....	22
Tabelle 8: Unterschiede in den Befundungszeiten in der Benignengruppe in Abhängigkeit vom Befunder sowie von der Anwendung des CAD-Systems:	22
Tabelle 9: Unterschiede in den Befundungszeiten in der Malignengruppe in Abhängigkeit vom Befunder sowie von der Anwendung des CAD-Systems:	23
Tabelle 10: Falsch-positive Markierungen in der Früherkennungsgruppe (n = 103): Ebenenabhängig, im Seitenvergleich und getrennt für Herdbefunde und Mikrokalk:	24
Tabelle 11: Falsch-positive Markierungen in der Benignengruppe (n = 102): Ebenenabhängig im Seitenvergleich und getrennt für Herdbefunde und Mikrokalk:	25
Tabelle 12: Falsch-positive Markierungen in der Malignengruppe (n = 98): Ebenenabhängig im Seitenvergleich und getrennt für Herdbefunde und Mikrokalk:	25
Tabelle 13: Richtig-positive Markierungen in der Malignengruppe (n = 98): Ebenenabhängig im Seitenvergleich und getrennt für Herdbefunde und Mikrokalk:	26

Abbildungsverzeichnis

Abbildung 1: Empirische Dichten und ROC-Kurven: kranke Patientinnen (rote Balken) und gesunde (blaue Balken). (a) Standardbeispiel, (b) Extremfall 1: nicht besser als der Zufall, (c) Extremfall 2: perfekte Trennung.	5
Abbildung 2: Senographie 2000D, Firma GE, Medical Systems, Buc, Frankreich.	7
Abbildung 3: Mammographie Workstation mit zwei hochauflösenden Monitoren (2000×2500 Pixel) für die Befundung. Zusätzlich abgebildet ist ein Computer zur Bedienung des radiologischen Informationssystems.	8
Abbildung 4: Struktur von clustered-data, x steht für eine Beobachtung (Werner 2006).	10
Abbildung 5: Altersverteilung der Früherkennungsgruppe.	12
Abbildung 6: Altersverteilung der Benignengruppe.	13
Abbildung 7: Histologie der Benignengruppe (A= Adenom; DH= duktale Hyperplasie; FI = Fibroadenom; AD = Adenose, PA = Papillom; AH = atypische duktale Hyperplasie; Z = Zyste).	13
Abbildung 8: Altersverteilung der Malignengruppe.	14
Abbildung 9: Histologie der Malignengruppe.	14
Abbildung 10: ROC-Kurven aller drei Befunder für die Malignen- und Früherkennungsgruppe ohne (links) und mit (rechts) Gebrauch des CAD-Systems.	15
Abbildung 11: Struktur von clustered-data, x steht für eine Beobachtungen.	17
Abbildung 12: ROC-Kurven aller drei Befunder für die Malignengruppe ohne (links) und mit (rechts) Gebrauch des CAD-Systems.	18
Abbildung 13: Box-Plots der Befundungszeiten in Abhängigkeit von der Patientengruppe, vom Befunder sowie von der Anwendung des CAD-Systems.	21
Abbildung 14: Darstellung der rechten Brust einer 66-jährigen Patientin aus der Benignengruppe. (A) In der MLO-Ebene drei falsch-positive Markierungen für Mikrokalk. (B) Vergrößerte Darstellung der drei falsch-positiven vaskulären Mikrokalkzifikationen.	27
Abbildung 15: Darstellung der rechten Brust einer 51-jährigen Patientin aus der Früherkennungsgruppe. (A) In der MLO-Ebene eine falsch-positive Markierung für Mikrokalk. (B) Vergrößerte Darstellung.	28
Abbildung 16: 65 Jahre alte Patientin aus der Früherkennungsgruppe. (A) In der MLO-Ebene der rechten Brust wurden keine Markierungen gesetzt hingegen eine falsch-positive Markierung für Herdbefund in der linken Brust (B). (C) Darstellung der entsprechenden linken MLO-Ebene ohne CAD-Markierungen. In der CC-Ebene der rechten Brust (D) und der linken Brust (E,F) wurden keine Markierungen gesetzt.	29
Abbildung 17: 55 Jahre alte Patientin mit histologisch gesichertem invasiv-duktalem Karzinom in der rechten Brust. (A) Richtig-positiv markierter Mikrokalk in der MLO-Ebene. (B) Vergrößerte Darstellung. (C) Richtig-positiv markierter Mikrokalk in der CC-Ebene, ebenfalls mit vergrößerter Darstellung (D). Zusätzlich eine falsch-positive Markierung eines Herdbefundes anhand der CC-Ebene.	30
Abbildung 18: 55 Jahre alte Patientin mit histologisch gesichertem invasiv-duktalem Karzinom in der rechten Brust. (A) In der rechten MLO-Ebene ist eine richtig-positive Markierung für Mikrokalk gesetzt worden. (B) Vergrößerte Darstellung der Region mit Mikrokalk. (C) In der rechten CC-Ebene wurden zwei richtig-positive Markierungen für Mikrokalk gesetzt. In der CC-Ebene wurde eine richtig-positive Markierung für Herdbefunde gesetzt. (D) Vergrößerte Darstellung der Region mit Mikrokalk.	31
Abbildung 19: 68 Jahre alte Patientin mit positivem Tastbefund links außen und histologisch gesichertem invasiv-duktalem Karzinom. (A) In der rechten MLO-Ebene wurden keine Markierungen gesetzt. (B) In der linken MLO-Ebene wurde eine richtig- positive Markierung für Herdbefund gesetzt. (C) Im Vergleich die linke MLO-Ebene ohne CAD-Markierung. In der CC-Ebene eine falsch-positive Markierung in der rechten Brust (D) und in der linken	

Brust (E). Ebenfalls wurde eine richtig-positive Markierung für einen suspekten Herdbefund in der linken Brust gesetzt (E). Entsprechende CC-Ebene der linken Brust ohne CAD-Markierungen.	32
Abbildung 20: 66 Jahre alte Patientin aus der Früherkennungsgruppe. (A) In der MLO-Ebene drei falsch positive Markierungen für Mikrokalk. (B) Vergrößerte Darstellung der polymorphen Mikrokalzifikationen (BIRADS 4). Histologie: Duktale Hyperplasie mit Mikrokalzifikationen.	33
Abbildung 21: Darstellung der linken Brust einer 65-jährigen Patientin aus der Benignengruppe. (A) In der linken CC-Ebene drei falsch positive Markierungen für Mikrokalk und eine falsch positive Markierung für Herdbefund. (B) Vergrößerte Darstellung.	34

Abkürzungsverzeichnis

A	Adenom
ACR	American College of Radiology
AD	Adenose
AH	Atypische duktale Hyperplasie
AUC	Area under the curve
BI-RADS	Breast Imaging Reporting and Data System
CAD	computerassistierte Diagnose
CC	kraniokaudal
DCIS	duktales Carcinoma in situ
DCIS + EIC	duktales Carcinoma in situ + extensiv-intraduktale Komponente
DH	duktales Hyperplasie
EUSOMA	European Society of Mastology
FI	Fibroadenom
IDC	invasiv-duktales Carcinoma in situ
IDC + DCIS	invasiv-duktales Karzinom + duktales Carcinoma in situ
IDC + EIC	invasiv duktales Karzinom + extensiv intraduktale Komponente
ILC	invasiv lobuläres Karzinoma
IPC	invasiv papilläres Karzinom
LCIS	lobuläres Carcinoma in situ
MEC	medulläres Karzinom
MLO	mediolateral oblique
Mo	Molybdän
PA	Papillom
Pat	Patientin
Rh	Rhodium
ROC	Receiving Operating Characteristics
TC	tubuläres Karzinom
Z	Zyste

Danksagung

Herrn Professor Dr. med. E. Grabbe danke ich für die Vergabe dieses Themas und die Möglichkeit zur Promotion in der Abteilung Diagnostische Radiologie der Medizinischen Fakultät der Georg-August-Universität Göttingen.

Mein besonderer Dank gilt Frau PD Dr. med. Silvia Obenauer für die engagierte Betreuung und Organisation bei der Erstellung dieser Dissertation.

Herrn Professor E. Brunner, Dr. Frank Konietschke und Dr. Carola Werner von der Abteilung Medizinische Statistik der Georg-August-Universität Göttingen danke ich für die Hilfestellung und Beratung bei der statistischen Auswertung. Des Weiteren danke ich Dr. Frank Konietschke für das Korrekturlesen dieser Arbeit.

Lebenslauf

Von Besim Cetin Angic

Ich wurde am 6. Februar 1976 in Böblingen geboren. 1982 wurde ich in die Ludwig-Uhland Grundschule in Böblingen eingeschult. Ab 1986 besuchte ich das Max-Planck-Gymnasium in Böblingen und legte dort 1996 meine Abiturprüfung ab. Im November 1996 nahm ich das Studium der Humanmedizin an der Georg-August-Universität in Göttingen auf. Die ärztliche Vorprüfung legte ich im Frühjahr 2003 ab. Im Laufe des Studiums famulierte ich in den Fächern Hals-Nasen-Ohrenheilkunde (Klinikzentrum Dogruk, Bursa, Türkei), Augenheilkunde (Universitätsklinik, Denizli, Türkei), Innere Medizin (Dipl.med Glückermann, Halberstadt), Chirurgie, (Universitätsklinik Denizli, Türkei). Mein Praktisches Jahr absolvierte ich in der Ammerlandklinik in Westerstede (Wahlfach Anästhesie). Mein Studium schloss ich mit dem 2. Staatsexamen im Juni 2007 ab. Seit dem 1. Oktober 2007 arbeite ich als Assistenzarzt in der Abteilung Augenheilkunde der Ammerlandklinik.