

**VERGLEICH DER DIAGNOSTISCHEN GENAUIGKEIT VON  
KONVENTIONELLER RÖNTGENMAMMOGRAPHIE MIT  
DER MR-MAMMOGRAPHIE IM LANGZEITVERLAUF**

**DISSERTATION**

**Zur Erlangung des akademischen Grades**

*doctor medicinae (Dr. med.)*

**Vorgelegt dem Rat der medizinischen Fakultät der  
Friedrich-Schiller-Universität Jena**

**von Matthias Benndorf**

**geboren am 18. August 1985 in Halle**

## **Gutachter**

- 1. Gutachter: Herr Prof. Dr. Werner A. Kaiser, Universitätsklinikum Jena**
- 2. Gutachter: Herr Prof. Dr. Ingo B. Runnebaum, Universitätsklinikum Jena**
- 3. Gutachter: Frau Prof. Dr. Christiane K. Kuhl, Universitätsklinikum Aachen**

*Tag der Einreichung: 03.02.2009*

*Tag der öffentlichen Verteidigung: 07.06.2010*

# Inhaltsverzeichnis

<b>1</b>	<b>Abkürzungsverzeichnis</b>	<b>1</b>
<b>2</b>	<b>Zusammenfassung</b>	<b>2</b>
<b>3</b>	<b>Einleitung</b>	<b>4</b>
3.1	Grundlegendes zur Epidemiologie maligner Neoplasien der Mamma .	4
3.2	Arten der Prävention . . . . .	4
3.2.1	Funktionsweise der sekundären Prävention . . . . .	5
3.2.2	Die Bedeutung falsch positiver und falsch negativer Befunde im Screening . . . . .	5
3.3	Tumorbiologie . . . . .	6
3.3.1	Tumorklassifikation . . . . .	8
3.4	Das angewandte Screeningverfahren . . . . .	9
3.4.1	Bildinformationen der Mammographie . . . . .	10
3.4.2	Bildinformationen der MR-Mammographie . . . . .	12
<b>4</b>	<b>Ziele der Arbeit</b>	<b>13</b>
<b>5</b>	<b>Methodik</b>	<b>14</b>
5.1	Studienaufbau . . . . .	14
5.1.1	Studienpopulation . . . . .	14
5.1.2	Studienaufbau . . . . .	14
5.1.3	Umgang mit dem Loss-To-Follow-Up . . . . .	15
5.1.4	Definition der falsch-negativen MR-Untersuchung . . . . .	15
5.2	Bewertung medizinischer Befunde . . . . .	15
5.2.1	Läsionsspezifische Kontrastmittelanreicherung . . . . .	17
5.2.2	Unterscheidung von mass- und non-mass-Tumoren . . . . .	20
5.3	Bildgebende Verfahren . . . . .	21
5.3.1	Mamma-Sonographie . . . . .	21
5.3.2	Konventionelle Röntgenmammographie . . . . .	21
5.3.3	MR-Mammographie . . . . .	22
5.3.4	Grund für die MR-Untersuchung . . . . .	23
5.4	Statistische Analyse . . . . .	23
5.4.1	Beurteilung der Güte diagnostischer Verfahren . . . . .	23
5.4.2	Kappa-Wert . . . . .	26

5.4.3	Chi-Quadrat-Test . . . . .	26
5.4.4	Mann-Whitney U Test . . . . .	27
5.4.5	Konfidenzintervalle für Korrelationskoeffizienten . . . . .	27
<b>6</b>	<b>Ergebnisse</b>	<b>29</b>
6.1	Ablauf der Studie . . . . .	29
6.1.1	Zeitlicher Ablauf . . . . .	29
6.1.2	Anteil ermittelter Röntgenmammographien und Follow-Up-Untersuchungen . . . . .	29
6.2	Eigenschaften der diagnostizierten Malignome . . . . .	30
6.3	Diagnostische Genauigkeit der Röntgenmammographie . . . . .	30
6.3.1	Problematik der Parenchymdichte . . . . .	31
6.4	Diagnostische Genauigkeit der MR-Mammographie . . . . .	34
6.5	Abweichende Befunde in Röntgenmammographie und MRM . . . . .	35
6.6	Malignomentwicklung und falsch negative Befunde . . . . .	36
6.6.1	Werdegang detektierter Malignome . . . . .	36
6.6.2	Falsch negative MRM-Befunde . . . . .	37
6.7	Die Messung des Tumordurchmessers . . . . .	37
6.7.1	Angaben zu Multifokalität und Multizentrität . . . . .	39
6.8	Die Reevaluation der MR-Mammographien . . . . .	39
6.8.1	Diagnostische Genauigkeit der Reevaluation . . . . .	39
6.8.2	Beschreibung der reevaluierten Herdbefunde . . . . .	41
6.8.3	Falsch negative MR-Befunde . . . . .	42
6.8.4	Gründe für falsch positive Befunde in der MRM . . . . .	42
<b>7</b>	<b>Diskussion</b>	<b>48</b>
7.1	Statistische Vorüberlegungen . . . . .	48
7.2	Einordnung der ermittelten diagnostischen Genauigkeiten . . . . .	49
7.2.1	Konventionelle Röntgenmammographie . . . . .	49
7.2.2	MR-Mammographie . . . . .	52
7.2.3	Der direkte Vergleich beider Bildgebungsmodalitäten . . . . .	54
7.3	Intervallkarzinome und falsch negative Befunde in der MRM . . . . .	55
7.4	Genauigkeit der Bildgebung hinsichtlich der Tumorausbreitung . . . . .	56
7.4.1	Tumordurchmesser . . . . .	56
7.4.2	Multifokalität und Multizentrität . . . . .	57
7.5	Die Reevaluation der MR-Mammographien . . . . .	59

7.5.1	Veränderte diagnostische Genauigkeit . . . . .	59
7.5.2	Objektivierte Gründe für falsch positive MR-Befunde . . . . .	60
<b>8</b>	<b>Schlussfolgerungen</b>	<b>62</b>
	<b>Literaturverzeichnis</b>	<b>64</b>
<b>9</b>	<b>Anhang</b>	<b>75</b>
9.1	Danksagung . . . . .	75
9.2	Ehrenwörtliche Erklärung . . . . .	76

# 1 Abkürzungsverzeichnis

**Tab. 1:** IN DER VORLIEGENDEN ARBEIT VERWENDETE ABKÜRZUNGEN

¬	logischer Negator	MRM	Magnet-Resonanz-Mammographie
∧	logisches „und“	MRT	Magnet-Resonanz-Tomographie
∨	logisches „oder“	NPW	negativer prädiktiver Wert
Abb.	Abbildung	PPW	positiver prädiktiver Wert
ACR	<i>American College of Radiology</i>	ROC	<i>receiver operating characteristic</i>
BIRADS	<i>breast imaging reporting and data system</i>	T2w-TSE	<i>T2-weighted-turbo-spin-echo</i>
DCIS	<i>duktales-carcinoma-in-situ</i>	Tab.	Tabelle
FA	<i>flip angle</i>	TE	Echozeit
FSU	Friedrich-Schiller-Universität	TIRM	<i>turbo-inversion-recovery-magnitude</i>
Gd-DTPA	<i>Gadolinium-diethylenetriamine-pentaacid</i>	TR	Relaxationszeit
IDC	invasiv duktales Karzinom		
ILC	invasiv lobuläres Karzinom		
LCIS	<i>lobuläres-carcinoma-in-situ</i>		
mm	Millimeter		
MR	Magnet-Resonanz		

## 2 Zusammenfassung

Die MR-Mammographie (MRM) gilt aktuell als das sensitivste Verfahren zur Detektion maligner Neoplasien der Mamma. Die Sensitivität wird zwischen 80 % und 100 % angegeben. Umstritten ist hingegen die Spezifität der Methode, hier variieren die Angaben enorm (65 % bis 97 %). Aufgrund der hohen Lebenszeitprävalenz maligner Neoplasien der Mamma (jede zehnte Frau erkrankt in der Bundesrepublik Deutschland an Brustkrebs) ist die Implementierung eines klinisch möglichst genauen Diagnoseverfahrens anzustreben. Die einzige derzeit anerkannte Screeningmethode ist die konventionelle Röntgenmammographie. Argumentiert wird für dieses Vorgehen mit der niedrigen Spezifität der MRM, der nur eingeschränkten Verfügbarkeit sowie den relativ hohen Kosten der Methode. Außerdem wird der nicht einheitliche Ausbildungsstand der Befunder kritisiert.

Das Ziel der vorliegenden Arbeit ist die Überprüfung der These, ob es einen Unterschied in der diagnostischen Genauigkeit von MRM und konventioneller Röntgenmammographie gibt. Desweiteren sollen objektivierbare Gründe für falsch positive MR-Befunde gesucht werden, die für eine eventuelle niedrige Spezifität der Methode verantwortlich sein könnten. Ebenfalls sollen falsch negative Befunde näher klassifiziert werden.

Die bisher publizierten Studien zu einer sinnvollen Angabe der diagnostischen Genauigkeit der MRM rekrutierten ihr Patientenkollektiv jeweils aus genetisch prädisponierten Hochrisikopatienten. Die Spezifität hängt von dem Anteil richtig negativer Untersuchungen ab. Um die klinische Situation repräsentativ abzubilden, ist die Verifikation derselben durch Betrachtung von Follow-Up-Untersuchungen unabdingbar. Eine Vorauswahl nur von histopathologisch verifizierten Befunden vermindert den Anteil richtig negativer Untersuchungen.

Die vorliegende Arbeit wurde als retrospektive Kohortenstudie angelegt. Die Studienkohorte rekrutierte sich konsekutiv aus der klinischen Anwendung der MRM, sie ist nach Kenntnis des Autors die erste Arbeit diesen Typs. Erfasst wurden die Befunde aller 342 Patienten, die im Zeitraum vom 01.01.2005 bis zum 20.03.2006 eine MRM an der FSU Jena erhielten und für diese eine Überweisung aus der Klinik für Frauenheilkunde und Geburtshilfe der FSU Jena hatten. Um den direkten Vergleich der Methode mit der Röntgenmammographie anzustellen, wurden die Befunde der zeitlich vor der MRM durchgeführten Röntgenmammographie erfasst (74,6 % der Patienten verfügten über eine solche). In 91,6 % der Fälle konnte eine Follow-Up-Untersuchung

ermittelt werden. Damit sind die ermittelten Schlussfolgerungen als statistisch valide anzusehen. Bei der Ermittlung der Follow-Ups wurde in 19,6 % der Fälle von extern Befunde angefordert. Als hinreichend wurden für das Follow-Up eine Röntgenmammographie, MRM, Ultraschalluntersuchung, histopathologische Absicherung oder eine klinische Untersuchung angesehen. In einem zweiten Arbeitsschritt wurden alle histopathologisch gesicherten intramammären Herdbefunde ( $n = 216$ ) von zwei Untersuchern reevaluiert und nach den aktuellen BIRADS-Kriterien bewertet.

Für die MRM (Röntgenmammographie) resultierten die folgenden statistischen Kennzahlen: Sensitivität 96,9 % (64,8 %), Spezifität 94,0 % (93,0 %), PPW 79,1 % (70,8 %), NPW 99,3 % (90,9 %) sowie diagnostische Genauigkeit 94,5 % (87,1 %). Statistisch signifikant war die MRM der Röntgenmammographie hinsichtlich ihrer Sensitivität, diagnostischen Genauigkeit und ihres NPW überlegen. Im Unterschied zur MRM hing die Genauigkeit der Röntgenmammographie stark von der Dichte des Brustdrüsenparenchyms ab. Vier Malignome wurden in der initialen MRM nicht detektiert. Die Reevaluation konnte in diesen Untersuchungen ebenfalls keine suspekten Läsionen ausmachen. In drei Fällen handelte es sich um Malignome mit einem Durchmesser von 2 mm oder weniger. Schon aufgrund biologischer Erwägungen erhebt die MRM hier keinen Anspruch auf eine genaue Diagnostik. Auffallend bei den falsch positiven Befunden der MRM war der überproportional hohe Anteil von non-mass-Läsionen, welche hier dreimal häufiger als unter allen evaluierten Herdbefunden vorkamen. Das Wachstumsmuster von non-mass-Läsionen ist weitaus schlechter untersucht als das typische der solide wachsenden Neoplasie. Eine strikte Befolgung der publizierten Empfehlungen hinsichtlich einer glatten Tumorberandung im Sinne von Gutartigkeit hätte einen Anteil der falsch positiven Befunde verhindern können. 58,8 % der falsch positiven Befunde waren glatt berandet (im Vergleich: 3,8 % der Malignome). Die MRM war sensitiver für multifokales und -zentrisches Tumorstadium, jedoch weniger spezifisch.

Insgesamt war die MRM der Röntgenmammographie überlegen. Sie führte zu einer ähnlich niedrigen Anzahl falsch positiver und einer deutlich geringeren Anzahl falsch negativer Befunde. Insbesondere bei jungen Patienten mit dichtem Brustdrüsenparenchym erscheint die hier insuffiziente Röntgenmammographie als alleiniges Diagnoseverfahren zweifelhaft. Weitere Studien mit dem Ziel einer verbesserten Differenzierung maligner von benignen non-mass-Läsionen sind notwendig, um den Anteil der falsch positiven Befunde der MRM weiter zu senken.

## 3 Einleitung

### 3.1 Grundlegendes zur Epidemiologie maligner Neoplasien der Mamma

In der Bundesrepublik Deutschland erkranken jährlich rund 57.000 Frauen an Brustkrebs. Damit sind maligne Neoplasien der Mamma die häufigste Krebserkrankung der Frau, sie machen hier nach aktuellen Erhebungen 27,8 % aller Krebsneuerkrankungen aus. Die Fünfjahres-Überlebensrate beträgt für dieses Patientenkollektiv 81 %. Die Inzidenz maligner Erkrankungen der Mamma ist stark altersabhängig, ab dem 35. Lebensjahr (Inzidenz: 53/100.000) steigt sie sprunghaft an und hat ihren Höhepunkt in der Altersgruppe der 65- bis 69-jährigen (Inzidenz: 303/100.000) (Bertz et al. 2008).

Anders ausgedrückt bedeuten diese Zahlen folgendes: Von allen 2003 in der Bundesrepublik Deutschland an Brustkrebs erkrankten Frauen erlebten krankheitsbedingt 10.830 das Jahr 2009 nicht. Absolut gesehen versterben in der Bundesrepublik pro Jahr etwa 17.000 Frauen an Brustkrebs. Hieraus wird klar, dass die Thematik Brustkrebs ein hohes wissenschaftliches und gesundheitspolitisches Interesse verdient.

### 3.2 Arten der Prävention

Ein erklärtes Ziel der modernen Medizin ist es, die Mortalität maligner Neoplasien der Mamma weiter zu senken. Um schlüssige Konzepte für diese Aufgabe zu implementieren, ist zunächst die Frage zu beantworten, auf welche Weise dies erreicht werden kann.

Eine ganze Reihe von Risikofaktoren wurden beschrieben, die die Entstehung eines Malignoms in der Brust begünstigen. Im Einzelnen sind dies eine familiäre Häufung von Malignomen in den Mammæ, der Lebensstil (Übergewicht, Bewegungsmangel, Ernährung), diverse endogene und exogene endokrine Faktoren sowie als stärkster Risikofaktor ein fortgeschrittenes Lebensalter (Albert und Schreer 2008). Im Falle der Lebensgewohnheiten sowie exogener endokriner Faktoren ist es möglich, Primärprävention zu betreiben. Unter Primärprävention versteht man die Gesundheitsförderung, bevor überhaupt eine Krankheit aufgetreten ist. Es ist also möglich, die Inzidenz des Brustkrebses durch eine konsequent in der Bevölkerung umgesetzte Änderung der Lebens- und Ernährungsgewohnheiten zu senken (Monninkhof et al.

2007).

Weiterhin besteht die Möglichkeit der Sekundärprävention, das heißt der Entdeckung von Krankheit in einem frühen Stadium. Die Sekundärprävention ist Aufgabe des Screenings. Screening bezeichnet die Untersuchung von asymptomatischen Individuen ohne vorbeschriebene Abnormitäten in dem zu untersuchenden Organ (Elmore et al. 1998).

### **3.2.1 Funktionsweise der sekundären Prävention**

Das Ziel der sekundären Prävention ist die Diagnose von Erkrankungen in einem frühen, noch gut therapierbaren Stadium. Die generelle Aussage aber, je früher eine Krankheit im Screening diagnostiziert wird, desto erfolgreicher das Screening, ist nicht zulässig.

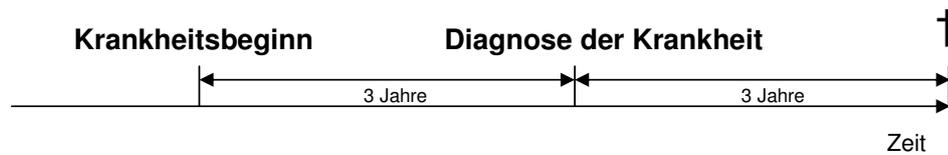
In die Literatur ist dieser potentielle Trugschluss unter dem Terminus Lead-Time-Bias eingegangen. So ist es möglich, bestimmte Erkrankungen mittels eines Screenings zwar zu detektieren und die scheinbare Überlebenszeit seit Diagnosestellung zu verlängern, die Gesamtüberlebenszeit jedoch nicht zu beeinträchtigen (Gates 2001). Abbildung 1 veranschaulicht diesen Sachverhalt.

Im Falle maligner Neoplasien der Mamma ist der tatsächliche Nutzen von Screeninguntersuchungen belegt. Ein konsequentes Screening führt demnach zu einer Senkung der Mortalität durch Brustkrebs. Für 50- bis 74-jährige Screeningteilnehmer beträgt das Relative Risiko an Brustkrebs zu versterben 0,74 im Vergleich zu Nicht-Screeningteilnehmern (Kerlikowske et al. 1995, Michaelson et al. 2002).

Neben der Lead-Time-Bias existiert eine weitere Störgröße von Screeninguntersuchungen, die Length-Time-Bias. In einem Screening werden überproportional häufig langsam wachsende, biologisch nur wenig aggressive Tumoren diagnostiziert. Grund hierfür ist das extrem heterogene Wachstumsverhalten von, im Falle dieser Arbeit relevanten, Mammakarzinomen, siehe 3.3. In einigen Fällen expandieren diese deutlich schneller als dass die bisher vorherrschenden Screeningmethoden eine Möglichkeit hätten, sie statistisch gesehen in einem frühen Stadium zu detektieren (Gates 2001).

### **3.2.2 Die Bedeutung falsch positiver und falsch negativer Befunde im Screening**

So sensitiv und spezifisch die verwendete Screeningmethode auch immer sein mag, es wird mit einem diagnostischen Test stets zu einer bestimmten Anzahl falsch negativer sowie falsch positiver Befunde kommen, siehe Tabelle 9 auf Seite 25. Aus der

**Ohne Screening****Mit Screening**

**Abb. 1:** Gezeigt ist der Effekt der sogenannten Lead-Time-Bias. Es wird eine fatale Erkrankung angenommen, die drei Jahre nach ihrer Diagnose (die im ersten Fall nicht durch Screening sondern klinisch gestellt wurde) zum Tod führt. Durch die Implementierung eines Screenings im zweiten Fall lässt sich der biologische Krankheitsbeginn auf zellulärer Ebene selbstverständlich nicht ändern. Ein Screening, welches die Diagnose um ein Jahr vorzieht, die Überlebenszeit nach Diagnosestellung jedoch lediglich um ein Jahr verlängert, verfehlt sein Ziel (i. e. die Senkung der Mortalität).

Definition des Screenings (siehe 3.2) wird ersichtlich, dass abhängig von der Prävalenz der entsprechenden Erkrankung nur wenige Screeningteilnehmer tatsächlich von diesem profitieren. Jedoch sind alle Teilnehmer dem Risiko eines falsch positiven Befundes ausgesetzt.

Falsch positive Befunde führen zu weiterer Diagnostik und lösen in dem Patienten unnötige Angst aus. Umgekehrt wiegen falsch negative Befunde den Teilnehmer in Sicherheit und können dazu führen, dass er klinische Symptome der progredienten Erkrankung in voller Überzeugung seiner Gesundheit ignoriert. Der Wert einer Screeningmethode wird daher maßgeblich durch ihre Sensitivität und Spezifität bestimmt (Cochrane und Holland 1971).

### 3.3 Tumorbiologie

Für die erfolgreiche Diagnostik einer malignen Neoplasie ist ein Grundverständnis ihres Wachstumsverhaltens notwendig. Mammakarzinome präsentieren sich weder mit einheitlicher Verdopplungszeit ihrer Größe noch einheitlicher Morphologie. Die Volumen-Verdopplungszeit schwankt bei Mammakarzinomen zwischen 98 und 951 Tagen, wobei hier das 10%-, respektive das 90%-Perzentil aller Tumoren angegeben

ist (von Fournier et al. 1994).

Tumorzellen verfügen über einen im Vergleich zu der nichtentarteten Ausgangszelle veränderten Stoffwechsel. So ermöglicht es die unphysiologische Überproduktion einer bestimmten Enzymklasse, der Matrix-Metallo-Proteinasen, dem Tumor aggressiv infiltrierend in das umliegende Gewebe hinein zu wachsen (Hojilla et al. 2008). Es besteht eine sich gegenseitig verstärkende Beziehung zwischen dem Wachstum von Tumorzellen und einer manifest werdenden perifokalen Entzündung (de Visser und Coussens 2006). Der Großteil aller malignen Tumoren der Brust löst, multifaktoriell bedingt, eine desmoplastische Reaktion des umliegenden Bindegewebes aus. Gemeint ist hier ein veränderter Umsatz der Extrazellulärmatrix im Sinne einer reaktiven Zunahme sowie der Retraktion fibrotischen Gewebes (Noel et al. 1993). In 30 - 50 % der malignen Neoplasien der Mamma kommt es zur Ablagerung von Calciumoxalat- und/oder Hydroxylapatitkristallen. Es ist gegenwärtig noch unklar, ob es sich dabei um eine Begleiterscheinung oder aber einen biologisch relevanten Mechanismus des Tumorwachstums handelt (Morgan et al. 2005).

In der Entwicklung einer malignen Neoplasie lassen sich weiterhin zwei unterschiedliche Wachstumsphasen unterscheiden. Bis zu einem Durchmesser von zwei bis drei Millimetern ist es dem Tumor möglich, ohne Neoangiogenese zu wachsen. Die Versorgung des neoplastischen Gewebes erfolgt hier durch bloße Diffusion. Der Durchmesser von drei Millimetern gilt in nicht-vaskularisierten Strukturen als kritische Grenze, an der ein Gleichgewicht zwischen Zellproliferation und -untergang erreicht wird (Holmgren et al. 1995). Grund hierfür sind zu lang werdende Diffusionswege.

Viele Neoplasien verweilen Monate bis Jahre in dieser Phase, ehe ein als *angiogenic switch* bezeichnetes Ereignis die Tumorzellen veranlasst, spezielle Wachstumsfaktoren wie bFGF (basic fibroblast growth factor) und VEGF (vascular endothelial growth factor) zu exprimieren. Ebenfalls werden angiogene Faktoren aus der Extrazellulärmatrix mobilisiert (Folkman 1995). Die diversen Wachstumsfaktoren führen zu Gefäßneubildungen.

Der sich weiter ausbreitende, hochgradig stoffwechselaktive Zellverband einer malignen Neoplasie hat einen enormen Bedarf an Nährstoffen. Diese werden ihm durch die neu einsprossenden Mikrogefäße zur Verfügung gestellt. Die Neoangiogenese eines Mammakarzinoms gilt als wichtiger prognostischer Faktor der Erkrankung. So korreliert die Dichte einsprossender Mikrogefäße mit dem Risiko der Metastasierung und der Wahrscheinlichkeit des Versterbens des Patienten (Weidner 1995).

Die Eigenschaften maligner Neoplasien sind nicht als qualitative Merkmale mit den binären Ausprägungen *vorhanden* oder *nicht vorhanden* zu verstehen. Vielmehr handelt es sich um quantitative Eigenschaften die in unterschiedlicher Ausprägung, abhängig von der Art der Neoplasie, vorkommen. Die Unterscheidung von Qualität und Quantität bei Krankheitserscheinungen ist ein bewährtes phänomenologisches Konzept (Canguilhem 1989).

### 3.3.1 Tumorklassifikation

Primäre Mammakarzinome entstammen dem ektodermalen Gewebeanteil der Brust, i. e. das Drüsenparenchym. Neben der Unterscheidung von invasiven und nicht-invasiven Formen folgt die weitere Beschreibung den anatomischen Gegebenheiten. Bedingt durch die Aufteilung der Brustdrüse in Ductus beziehungsweise Ductuli und Lobuli werden Mammakarzinome weiter entsprechend ihres Ursprungs in duktale und lobuläre Karzinome unterteilt, siehe Tabelle 2. Der Übergang von nicht-invasiven Karzinomen in eine invasive Form ist möglich (van Diest 1999).

**Tab. 2:** KLASSIFIKATION DER PRIMÄREN MAMMAKARZINOME NACH INVASIVITÄT UND ANATOMISCHEM URSPRUNG

Mammakarzinom			
nicht-invasiv*		invasiv	
DCIS	LCIS	IDC	ILC

\*nicht-invasiv bedeutet, dass das Karzinom die dem Drüsenepithel angrenzende Basalmembran als Wachstumsbeschränkung akzeptiert (carcinoma-in-situ). DCIS: duktales carcinoma-in-situ, LCIS: lobuläres carcinoma-in-situ, IDC: invasiv duktales Karzinom, ILC: invasiv lobuläres Karzinom

Neben den in Tabelle 2 gezeigten existieren weitere, seltenere Formen maligner Neoplasien der Mamma. Es handelt sich dabei um den Morbus Paget (Karzinom der intramammilär gelegenen Ductus), das inflammatorische Karzinom, medulläre Karzinome, muzinöse Karzinome sowie Neoplasien mesodermaler Gewebsanteile, i. e. Sarkome. Gutartige Tumore der Mamma sind Fibroadenome, Papillome und das Cystadenoma phylloides. Die häufigsten benignen Veränderungen der Mamma entstammen dem sogenannten mastopathischen Formenkreis. Bedingt durch ihre hohe Prävalenz sind sie verantwortlich für den Großteil aller intramammären Beschwerden (Santen und Mansel 2005). Daraus folgt, dass sie auch den absolut gesehen größ-

ten Anteil aller falsch positiven Befunde ausmachen, sowohl von der konventionellen Röntgenmammographie als auch der MRM.

Von klinischem Interesse ist auch das intramammäre Ausbreitungsmuster der Neoplasien. Man unterteilt dieses in ein monofokales, multifokales und multizentrisches Muster. Monofokal bedeutet, es existiert lediglich ein singulärer Primärtumor. Multifokalität steht für mindestens zwei separierte Herde, die auf einen Quadranten der Mamma beschränkt sein müssen. Multizentrität schließlich bezeichnet mindestens zwei Herde in unterschiedlichen Quadranten der Mamma (Bedrosian et al. 2003). Je disseminierter das intramammäre Tumorwachstum, desto schlechter ist die Langzeitprognose des Patienten, auch die nach einer abgeschlossenen Therapie (Kurtz et al. 1990).

### **3.4 Das angewandte Screeningverfahren**

Vorangestellt seien zwei als a priori gültig zu bezeichnende Aussagen. Erstens, ein medizinisches Screening sollte mittels der akkuratesten und gleichzeitig verfügbaren Methodik erfolgen. Zweitens, das kurzfristige und langfristige Wohl der Patienten, respektive Screeningteilnehmer, steht im Mittelpunkt der Untersuchung.

Die aktuell nach Leitlinien der *Planungskommission der Konzertierten Aktion Brustkrebs-Früherkennung in Deutschland* einzige allgemein wirksame Methode zur Brustkrebs-Früherkennung ist die konventionelle Röntgenmammographie (Albert et al. 2008). Deren Nutzen ist für Frauen ab dem 40. Lebensjahr belegt. Hier übersteigt erstmals der individuell aus dem Screening gezogene Nutzen das Risiko durch die Röntgenstrahlenbelastung (Nothacker et al. 2007). Bei dichtem Drüsengewebe oder aber einer familiären Disposition fordern die Leitlinien eine additive Mamma-Sonographie, um die Sensitivität der Röntgenmammographie anzuheben. Empfehlungen das zeitliche Intervall zwischen den einzelnen Mammographien betreffend schwanken zwischen einem und zwei Jahren (Fletcher und Elmore 2003).

Die Angaben für Sensitivität und Spezifität der Mammographie variieren. In einer Screeningstudie, die zusätzlich zu dem visuellen Urteil des Befunders ein CAD System (Computer Aided Detection, englisch für computergestützten Auswertungsmodus) zur Diagnostik verwendete, betrug die Sensitivität der Methode insgesamt 87,2 %, die Spezifität 96,6 % sowie der positive prädiktive Wert 18,0 % (Gilbert et al. 2008). Andere Studien geben Werte von 66 % und 45 % für die Sensitivität, respekti-

ve 75 % für die Spezifität der Röntgenmammographie an (Sardanelli et al. 2004, Berg et al. 2004). Maligne Neoplasien der Mamma haben eine Lebenszeitprävalenz von zehn Prozent. Die Screeningpopulation besteht aus allen asymptomatischen Frauen der Bundesrepublik Deutschland. In einer, zu einem bestimmten Zeitpunkt stattfindenden, Querschnittsuntersuchung dieser Screeningpopulation, werden mammographisch durchschnittlich 7,1 maligne Läsionen je 1000 Screeningteilnehmer diagnostiziert. Selbst in Anbetracht der eventuell niedrigen Sensitivität von 50 % wären also maximal 1,4 % der Screeningteilnehmer zu einem bestimmten Zeitpunkt an einem Malignom erkrankt (Gilbert et al. 2008).

### **3.4.1 Bildinformationen der Mammographie**

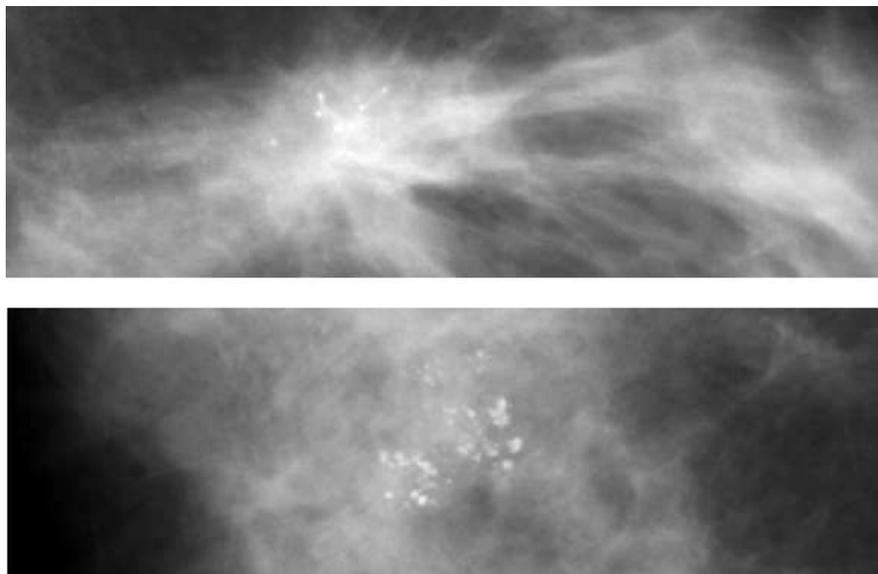
Die Mammographie ist für unterschiedliche Subpopulationen der Screeningpopulation unterschiedlich gut geeignet zur Detektion maligner Neoplasien. Ganz entscheidend hängt die Aussagekraft der Mammographie von der Zusammensetzung des Brustdrüsenparenchyms sowie des Stromas ab. Je dichter das Drüsenparenchym ist, desto niedriger ist die Sensitivität der Röntgenmammographie. Beträgt der Parenchymanteil innerhalb einer Mamma  $\geq 75\%$ , sinkt die Sensitivität auf unter 28 % (Pisano et al. 2008). Die Veränderung des Stromas im Sinne einer fibrozystischen Mastopathie verschlechtert die diagnostische Genauigkeit der Röntgenmammographie weiter. Tabelle 3 gibt eine Übersicht über die gängige Einteilung der intramammären Parenchymdichte.

Ein Befund kann sich auf drei Arten als malignomverdächtig im Mammogramm präsentieren. Erstens können suspekte, solide Herdbefunde, je nachdem in welchem Ausmaß sie Röntgenstrahlen absorbieren, in verschiedenen Graustufen zur Darstellung gebracht werden. Bewertet werden an diesen Herden Binnen- und Randkonfiguration. Verschiedene Deskriptoren lassen sich dem Malignitätsgrad einer Läsion zuordnen. So weist beispielsweise ein irregulär konfigurierter, spikulierter Herd auf ein Malignom hin (Fischer 2008). Zweitens treten bei einem gewissen Anteil der Malignome, jedoch auch der gutartigen Befunde, röntgendichte Mikroverkalkungen auf. Drittens kann ein Malignom einen Architekturdefekt im Parenchym verursachen, der jedoch nicht als Herdbefund abgrenzbar ist. Exemplarische Darstellungen finden sich in Abbildung 2.

**Tab. 3:** BRUSTPARENCHYMDICHTE NACH DEM ACR

ACR-Typ*	Beschreibung	Anteil des Drüsenparenchyms in %
1	Die Brust besteht fast vollständig aus Fettgewebe	$\leq 25$
2	Es finden sich verstreute fibroglanduläre Verdichtungen	25 - 50
3	Das Brustgewebe ist inhomogen dicht	51 - 75
4	Das Brustgewebe ist extrem dicht	$\geq 75$

\*ACR steht für American College of Radiology, nach welchem die gängige Einteilung der Parenchymdichte der Mamma angegeben wird, Tabelle nach (Saleh et al. 2005).



**Abb. 2:** Oben ist ein klassischer röntgendichter solider Tumor zu sehen, das spikulierte Wachstum weist stark auf die Malignität des Tumors hin. Im unteren Bild sind irreguläre, disseminierte Mikroverkalkungen zu sehen.

### **3.4.2 Bildinformationen der MR-Mammographie**

Die MR-Mammographie beurteilt die reine Morphologie von Läsionen in T1 und T2 gewichteten Sequenzen nach einem ähnlichen Schema wie die Röntgenmammographie. Mikrokalzifikationen sind in der MRT nicht darstellbar (Houserkova et al. 2008). Im Unterschied zu der konventionellen Mammographie liefert die MRM jedoch auch funktionelle Eigenschaften der Läsion, zu diesem Zweck wird dem Patienten intravenös ein Kontrastmittel appliziert. Gemäß den in 3.3 dargelegten Eigenschaften solider Herdbefunde ist eine zusätzliche Perfusion von soliden Strukturen ab einer gewissen Größe obligat. Das Kontrastmittel stellt so zum einen die intramammären und zum anderen die läsionsspezifischen Perfusionsverhältnisse dar (Kaiser und Zeitler 1989). Bestimmte Anreicherungsmuster erlauben die Unterscheidung von malignen und benignen Befunden, siehe Tabelle 6 auf Seite 19.

Die Angaben für die Sensitivität der MR-Mammographie variieren nur gering, sie bewegen sich zwischen 80 % und 100 %. Die Spezifität wird in aktuellen Studien mit der konventionellen Röntgenmammographie gleichgesetzt, oftmals werden jedoch auch deutlich niedrigere Werte angegeben. Die Angaben variieren hier zwischen 65 % und 97 % (Heywang-Köbrunner et al. 1997, Van Goethem et al. 2006, Kuhl 2007b).

## 4 Ziele der Arbeit

Die vorliegende Studie dient der Evaluation der diagnostischen Genauigkeit der MR-Mammographie sowie deren Vergleich mit der konventionellen Röntgenmammographie. Diagnostische Genauigkeit bezeichnet die Proportion tatsächlich zutreffender Befunde unter der Gesamtheit der Befunde. Um den direkten Vergleich der beiden Methoden anstellen zu können, ist es notwendig eine Kohorte von Patienten zu untersuchen, die sich, möglichst vollständig, diesen beiden diagnostischen Prozeduren zeitnah unterzieht. Die Validität der benignen Befunde soll durch die Erfassung von Follow-Up-Untersuchungen für den einzelnen Patienten gewährleistet werden.

Die Erfassung der diagnostischen Genauigkeit der jeweiligen Untersuchung im Langzeitverlauf ermöglicht es, Antworten zu folgenden Fragestellungen zu geben: Wieviele der als benigne bezeichneten Befunde müssen als falsch negativ angesehen werden? Wieviele Malignome entwickeln sich zwischen Röntgenmammographie respektive MR-Mammographie und dem Follow-Up in der Kohorte neu? Wie groß ist der Anteil an diskordanten Befunden in Mammographie und MR-Mammographie und welche Befunde sind im Langzeitverlauf akkurater?

Weiterhin sollen die Gründe für Fehlbefundungen in der MRM eruiert werden. Anlass hierfür ist die oftmals als ungenügend bezeichnete Spezifität der MR-Mammographie. Gibt es also objektivierbare, vermeidbare Gründe insbesondere für falsch positive Befunde? Auch lässt sich untersuchen, ein wie großer Anteil an Befunden durch eine Zweitevaluation hin verändert wird – sowohl zur richtigen als auch falschen Alternative hin.

Es können somit folgende zentrale Hypothesen aufgestellt werden, deren Überprüfung die vorliegende Studie gilt:

$H_0$ : Es gibt keinen Unterschied der diagnostischen Genauigkeit zwischen Mammographie und MR-Mammographie.

$H_1$ : Die MR-Mammographie hat eine der konventionellen Röntgenmammographie überlegene diagnostischen Genauigkeit.

## 5 Methodik

### 5.1 Studienaufbau

#### 5.1.1 Studienpopulation

Die Einschlusskriterien, nach denen im einzelnen die Patienten für die Studie ausgewählt wurden, waren die folgenden: Eine stattgefundene MR-Untersuchung der Mammae im Zeitraum vom 01.01.2005 bis zum 20.03.2006 in dem Institut für Diagnostische und Interventionelle Radiologie der FSU Jena aufgrund einer Überweisung für diese Untersuchung aus der Klinik für Frauenheilkunde und Geburtshilfe der FSU Jena sowie ein Eintrag im Aktenverzeichnis eben dieser Klinik mindestens ab dem Jahr 2005. Insgesamt wurden damit 342 Patienten in die Studie eingeschlossen.

Das mediane Alter der Patienten betrug zum Zeitpunkt der MR-Untersuchung 54,0 a, das mittlere Alter 54,4 a. Der Anteil männlicher Patienten belief sich in diesem Kollektiv auf 0,8 % ( $n = 3$ ).

#### 5.1.2 Studienaufbau

Die vorliegende Studie wurde als retrospektive Kohortenstudie angelegt. Die Kohorte rekrutierte sich dabei nach den oben aufgeführten Kriterien.

Insgesamt wurden für jeden Patienten die Ergebnisse von drei Untersuchungen vermerkt: Erstens einer konventionellen Röntgenmammographie vor der MR-Untersuchung im Zeitraum vom 01.01.2005 bis zum 20.03.2006. Zweitens die der MR-Mammographie als solcher und drittens einer Follow-Up-Untersuchung. Als hinreichend wurden für das Follow-Up befunden: eine Röntgenmammographie, eine MR-Mammographie, eine Mamma-Sonographie, ein klinisches Follow-Up oder eine histologische Absicherung. Als zeitlicher Abstand zwischen der Follow-Up-Untersuchung und der MR-Untersuchung wurden 24 Monate angestrebt.

Soweit vorhanden wurden initiale Röntgenaufnahmen und Follow-Ups verwendet, die an der FSU Jena angefertigt wurden oder aber solche, die in der jeweiligen Patientenakte hinterlegt waren. War dies nicht möglich, wurden die Befunde aus diversen Quellen angefordert, i. e. niedergelassene Ärzte und andere Kliniken als der FSU.

Im Falle einer histologischen Abklärung eines suspekten intramammären Herdes wurde das histopathologische Ergebnis vermerkt. Hierbei wurde auf Daten aus dem

Institut für Pathologie der FSU Jena sowie in zwei Einzelfällen auf auswärtige histopathologische Ergebnisse zurückgegriffen.

### 5.1.3 Umgang mit dem Loss-To-Follow-Up

Medizinische klinische Studien die ein Patientenkollektiv über einen definierten Zeitraum beobachten, können unter dem so genannten „Loss-To-Follow-Up Problem“ leiden. Dies bezeichnet den Fall, dass Patienten nach ihrer Aufnahme in die Studie keine Follow-Up-Untersuchungen mehr erhalten. Je geringer der Anteil an Patienten mit einer Follow-Up-Untersuchung ist, desto weniger valide sind die Aussagen, die anhand der Daten der verbleibenden Kohorte über die gesamte Kohorte getroffen werden können. (Sims 1973, Britton et al. 1995). Bereits 1973 empfahl Sims, eine maximale Rate von 10 % im Verlust von Follow-Ups zuzulassen.

### 5.1.4 Definition der falsch-negativen MR-Untersuchung

Um die Sensitivität und Spezifität eines diagnostischen Tests im Langzeitverlauf zu bestimmen, ist die Festlegung eines zeitlichen Grenzwertes für benigne Befunde notwendig. Wird nach dieser Zeitdauer eine Neoplasie diagnostiziert, wird die eigentliche Untersuchung nicht mehr als falsch negativ bewertet. Berücksichtigt werden muss hierbei das extrem heterogene Wachstumsverhalten maligner Neoplasien der Mammae.

So haben 10 % aller Mammatumoren eine Verdopplungszeit ihres Volumens von weniger als 100 Tagen, die mediane Verdopplungszeit liegt bei 260 Tagen (von Fournier et al. 1994). Aufgrund tumorbiologischer Erwägungen gelten Tumoren ab einer Größe von 3 mm als detektierbar in der MRT (Folkman 1995), siehe 3.3.

Ein Grenzwert von einem Jahr wurde in einer großangelegten Mammographie-Studie vorgeschlagen (Elmore et al. 1998). Die vorliegende Arbeit übernimmt unter Kenntnis der dargelegten Verhältnisse diese Festlegung. Unterschieden wurde weiterhin zwischen falsch negativen Befunden, die auf einen ärztlichen Fehler zurückzuführen waren, und solchen, die der Methode an sich zuzuschreiben waren.

## 5.2 Bewertung medizinischer Befunde

In einem ersten Schritt wurden die bereits erhobenen Befunde (i. e. Röntgenmammographien, MR-Mammographien sowie die Follow-Up-Untersuchungen) analysiert. Die jeweils eruierten Informationen sind in Tabelle 4 dargestellt.

Tab. 4: INFORMATIONSGEHALT DER EINZELNEN UNTERSUCHUNGEN

	Untersuchung		
	Mammographie	MR-Mammographie	Follow-Up
<b>Information</b>	Datum	Datum	Datum
	Herdanzahl	Herdanzahl	
	Durchmesser größter Herd	Durchmesser größter Herd	
	monofokal/ multifo- kal/ multizentrisch	monofokal/ multifo- kal/ multizentrisch	
	Drüsengewebsdichte nach ACR*		
<b>Gesamtbewertung</b>	nach BIRADS**	benigne, maligne, subioptisch vorher ge- sicher benigne oder maligne, histologische Abklärung erbeten	unauffällig oder su- spekt

\* Zur Erklärung der Einteilung der Parenchymdichte nach dem American College of Radiology (ACR) siehe Tabelle 3 auf Seite 11.

\*\* Zur Erklärung der BIRADS-Klassifikation siehe Tabelle 6 auf Seite 20.

Die histopathologische Einordnung umfasste folgende Merkmale: Die TNM Klassifikation des Tumors, im Falle von Malignität das Malignitätsgrading G1 bis G3, eine beschriebene Mono- oder Multifokalität, eine beschriebene Multizentrität, den Durchmesser des größten Herdes und die Art des Tumors.

In einem zweiten Schritt wurden sämtliche MR-Untersuchungen, nach oder vor denen eine histologische Abklärung suspekter und fraglicher Herde erfolgt war oder nach denen der Patient einen Tumor entwickelte, reevaluiert. Insgesamt handelte es sich dabei um 189 MR-Untersuchungen beider Mammae, von denen 216 Mammae von Interesse waren. Die Differenz wurde von dem in 27 Fällen bestehenden Verdacht einer bilateralen Erkrankung der Mammae verursacht. Zwei Beobachter (Ein Radiologe in Ausbildung mit einer Befundungserfahrung von ca. 5000 MR-Mammographien zum Zeitpunkt der Reevaluation) sowie der Verfasser dieser Arbeit (Befundungserfahrung von ca. 500 MR-Mammographien zum Zeitpunkt der Reevaluation) bewerteten dabei verblindet die MR-Mammographien erneut. Im Falle der Diskordanz wurde abklärend diskutiert, was jedoch nicht zwangsläufig zu einer Bewertungsänderung eines der Beobachter führen musste. Die Auswertung erfolgte unter

den auf Seite 22 beschriebenen Bedingungen. Die systematische Befundung suspekter Läsionen folgte dabei weitgehend den BIRADS Kriterien für MR-Mammographie (Ikeda et al. 2001), siehe Tabelle 5. Tabelle 6 gibt einen Überblick über die BIRADS-Klassen. Die Abbildungen 3 bis 5 bilden beispielhaft BIRADS-Deskriptoren ab, vergleiche ebenfalls die Abbildungen 7 und 8.

**Tab. 5:** MORPHOLOGISCHE BESCHREIBUNG VON KONTRASTMITTEL-ANREICHERNDEN LÄSIONEN IN DER MRT NACH BIRADS KRITERIEN

abnorme Kontrastmittelanreicherung*	
mass	non-mass
<b>Form:</b> rund, oval, lobuliert, irregulär	<b>räumliche Verteilung:</b> fokal, linear, duktal, segmental, regional, multipel regional, diffus
<b>Berandung:</b> glatt, irregulär, spikuliert	
<b>internes Anreicherungsmuster:</b> zentrifugal, zentripetal, homogen, heterogen, dunkle Septierungen, anreichernde Septierungen	<b>internes Anreicherungsmuster:</b> homogen, heterogen, stippled, clumped, retikulär-dendritisch

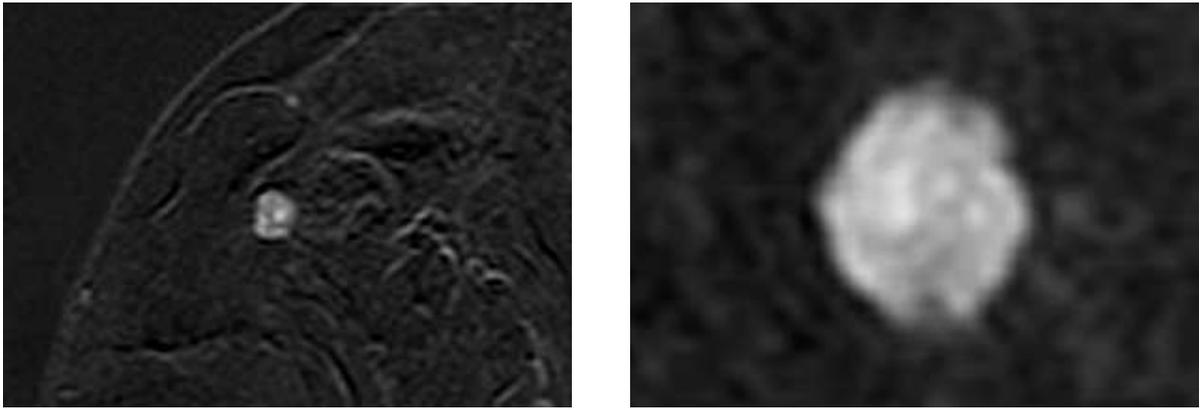
\* abnorme Kontrastmittelanreicherung ist definiert als Anreicherung mit höherer Signalintensität als das umliegende, normale Drüsengewebe. Tabelle nach (Tardivon et al. 2007). Zur Unterscheidung von mass- und non-mass-Läsionen siehe 5.2.2.

Die Klassifikation „foci-Läsionen unter 5 mm“ wurde in der vorliegenden Arbeit aus tumorbiologischen Erwägungen nicht übernommen. Solche Läsionen wurden anhand der obig dargestellten Kriterien beurteilt.

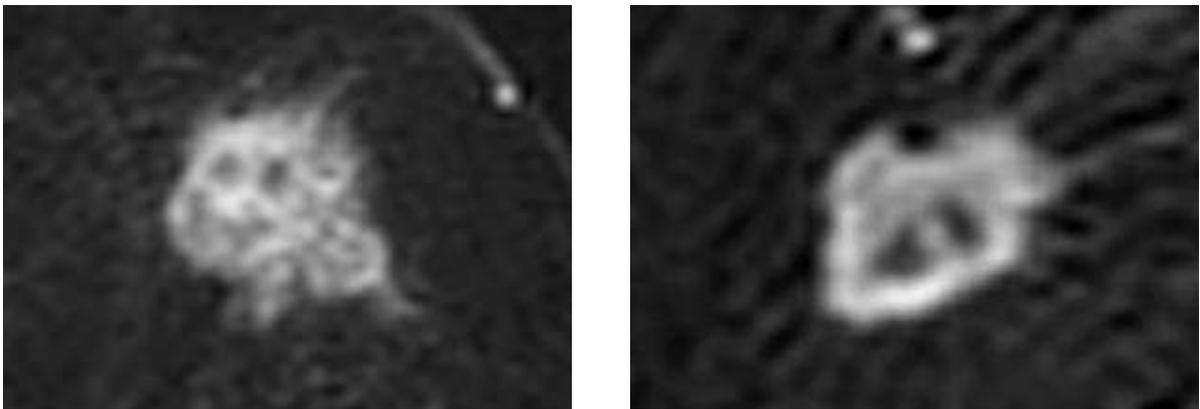
### 5.2.1 Läsionsspezifische Kontrastmittelanreicherung

Neben der Morphe der jeweiligen Struktur wurde auch das Anreicherungsmuster des applizierten Kontrastmittels in einer manuell gesetzten Region-of-Interest aufgenommen. Die Differenzierung maligner von benignen Läsionen beruht unter anderem auf den pharmakokinetischen Eigenschaften der Anreicherung des intravenös applizierten Kontrastmittels. Während ein kontinuierlicher Anstieg der Signalintensität ein starker Prädiktor für eine benigne Läsion ist, gilt das sogenannte Washout-Phänomen (Signalverlust über die Zeit nach einer initialen Anreicherung) als starker Prädiktor für eine maligne Läsion (Kaiser 1992, Boetes et al. 1994, Kuhl et al. 1999), siehe Abbildung 6.

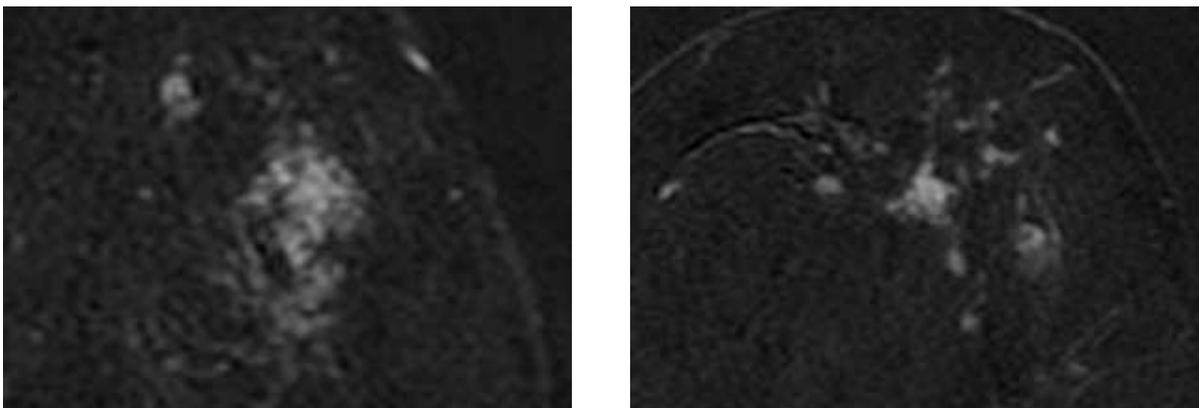
Ein Washout-Phänomen liegt vor, wenn der Herd 10% oder mehr seiner Signalintensität zwischen der ersten und siebten Minute nach der Kontrastmittelinjektion



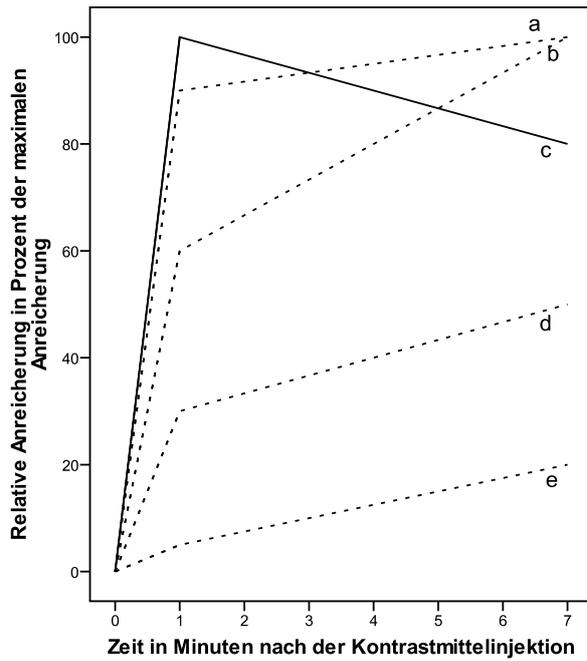
**Abb. 3:** Links ist eine glatt berandetes Fibroadenom zu sehen, rechts eine heterogen anreichernde, maligne mass-Läsion. Es handelt sich jeweils um Subtraktionsbilder.



**Abb. 4:** Links ist ein lobulierter, irregulär berandeter maligner Tumor zu sehen, der darüber hinaus über anreichernde innere Septen verfügt (*enhancing septations*). Rechts ist eine randständige Anreicherung zu sehen im Sinne einer zentripetalen Anreicherung. Es handelt sich jeweils um Subtraktionsbilder.



**Abb. 5:** Links ist eine non-mass-Läsion vom *stippled*-Typ zu sehen, vereinbar am ehesten mit einem benignen Befund. Rechts ist eine non-mass-Läsion vom *clumped*-Typ im Sinne eines DCIS dargestellt. Es handelt sich jeweils um Subtraktionsbilder.



**Abb. 6:** Dargestellt ist das Kontrastmittelanreicherungsverhalten benigner Herde im Sinne einer kontinuierlichen Anreicherung (Kurven *a*, *b*, *d*, *e*). Kurve *c* zeigt das Washout-Phänomen einer malignen Läsion entsprechend.



**Abb. 7:** Dargestellt sind die Subtraktionsbilder eines irregulär berandeten mass-Tumors, oben eine Minute nach Kontrastmittelinjektion, unten sieben. Beachte das zentrale Washout-Phänomen.

Tab. 6: DIE BIRADS-KLASSEN (RÖNTGENMAMMOGRAPHISCH)

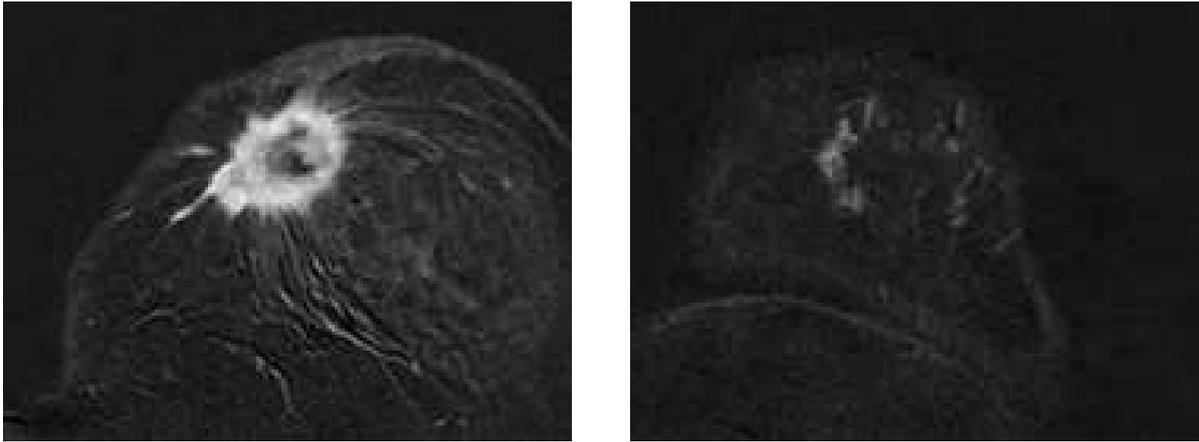
Beurteilungskategorie	Erklärung	Konsequenz
0	Beurteilung unvollständig, nicht möglich	weitere Bildgebung zur Gesamtbeurteilung erforderlich
1	kein Herdbefund	turnusgemäßes Screening
2	benigner Herdbefund	
3	wahrscheinlich gutartiger Befund, Malignomwahrscheinlichkeit $\leq 2\%$	Verlaufskontrolle in einem kurzen Intervall (6 Monate)
4	malignomsuspekter Befund	histologische Abklärung erforderlich
5	hochgradiger Verdacht auf ein Malignom	
6	histologisch bereits verifiziertes Malignom	

Tabelle nach (Saleh et al. 2005).

verliert. Umgekehrt gilt es als kontinuierliche Anreicherung, wenn der Herd von der ersten bis zur siebten Minute nach der Kontrastmittelinjektion 10 % oder mehr an Signalintensität hinzugewinnt (Jacobs et al. 2003). Werden sowohl diese obere als auch untere Grenze nicht überschritten, gilt die Kontrastmittelanreicherung als Plateauförmig. Eine geringe initiale Anreicherung (in der ersten Minute) spricht für die Gutartigkeit einer Läsion, vergleiche Kurven *b,d* und *e* in Abbildung 6.

### 5.2.2 Unterscheidung von mass- und non-mass-Tumoren

In der Reevaluation der MR-Untersuchungen wurden Tumoren unter anderem in die Klassen mass und non-mass eingeteilt, siehe Abbildung 8. Es handelt sich hierbei um Wachstumsmuster von Tumoren. Im Falle der mass-Tumoren wachsen diese das umliegende Gewebe mechanisch verdrängend. Non-mass-Tumoren hingegen respektieren anatomische Gegebenheiten und wachsen entlang dieser, sie verdrängen das umliegende Gewebe nicht. Aktuelle Publikationen beschäftigten sich mit der Problematik der akkuraten Diagnostik von intramammären non-mass-Tumoren (Tozaki und Fukuda 2006, Sakamoto et al. 2008).



**Abb. 8:** Links ist eine invasiv in das umliegende Gewebe wachsende mass-Läsion dargestellt (Deskriptoren: irregulärer Rand, spikuliert, zentripetale Kontrastmittelanreicherung), rechts eine sich diffus ausbreitende non-mass-Läsion. Beide Subtraktionsbilder wurden eine Minute nach Applikation des Kontrastmittels aufgenommen.

## 5.3 Bildgebende Verfahren

### 5.3.1 Mamma-Sonographie

Die vorliegende Studie stellt den direkten Vergleich zwischen dem Informationsgehalt der konventionellen Mammographie und der MR-Mammographie an. Ergebnisse der Mamma-Sonographie wurden auf zwei Arten in die Studie integriert. Erstens (bei fehlendem initialen Röntgenmammogramm oder aber einem unauffälligen Röntgenmammogramm und suspekter Sonographie) als Indikation zur MR-Untersuchung und zweitens als hinreichende Follow-Up-Methode. Da die verwendeten Sonographie-Ergebnisse aus den unterschiedlichsten Quellen stammten, wird auf die Darlegung einzelner technischer Spezifikationen verzichtet.

### 5.3.2 Konventionelle Röntgenmammographie

Da die verwendeten konventionellen Röntgenmammographien aus den verschiedensten Quellen stammten, wird an dieser Stelle auf die Aufführung sämtlicher technischer Spezifikationen einzelnen Aufnahmetechniken verzichtet. Allgemein lässt sich festhalten, dass bei einer Routinemammographie der Patient im Mittel einer Strahlungs-dosis von 3,8 mGy ausgesetzt wird. Im Standardprocedere werden von jeder Mamma eine Aufnahme in medio-lateral-obliquem und eine in cranio-caudalem Strahlengang angefertigt (Chevalier et al. 2004, Brnic und Hebrang 2001).

### 5.3.3 MR-Mammographie

Sämtliche MR-Untersuchungen erfolgten in einem Siemens SONATA MR-Tomographen oder einem Siemens MAGNETOM Symphony MR-Tomographen unter einer magnetischen Flussdichte von 1,5 Tesla. Das Untersuchungsprotokoll war Tabelle 7 entsprechend standardisiert. Es wurden je Messserie, mit Ausnahme der Lokalisierungssequenz, die lediglich der Überprüfung der optimalen Lagerung des Patienten diente, 33 transversale Aufnahmen beider Mammae angefertigt.

**Tab. 7:** MRM-UNTERSUCHUNGSPROTOKOLL

	Orientierung	TR in ms	TE in ms	FA in Grad	Schicht- dicke in mm	Schicht- anzahl	Voxelgröße in mm
1. Lokalisierung	Koronar/ sagital	15	5,15	40	10	4	$1,6 \times 1,6 \times 10$
2. T1 nativ	transversal	113	4,56	80	3	33	$1,1 \times 0,9 \times 3,0$
3. Dynamik	transversal	113	4,56	80	3	33	$1,1 \times 0,9 \times 3,0$
4. T2w-TSE	transversal	8900	191	90	3	33	$0,8 \times 0,7 \times 3,0$
5. TIRM	transversal	8420	70	180	3	33	$1,7 \times 1,4 \times 3,0$

T2w-TSE: T2-weighted Turbo-Spin-Echo, TR: Relaxationszeit, TE: Echozeit, FA: Flipwinkel (flip angle), Matrix: Auflösung in pixel, TIRM: Turbo Inversion Recovery Magnitude

Nach der Aufnahme der nativen, T1 gewichteten Bildserie wurde mittels eines maschinellen Injektors intravenös 0,1 mmol/(kg Körpergewicht) Gd-DTPA (Magnevist, Bayer HealthCare, Leverkusen, Deutschland) als Bolus von 3 ml/s, gefolgt von 20 ml NaCl, injiziert. Als Applikationsort wurde standardisiert eine Cubitalvene gewählt. Bei Gd-DTPA handelt es sich um ein Kontrastmittel, welches seit 1986 zur Darstellung perfundierter Strukturen in der dynamischen MR-Bildgebung der Mammae verwendet wird (Heywang et al. 1986, Kaiser und Zeitler 1989).

Die Aufnahmen der Kontrastmittelanreicherung erfolgten T1 gewichtet unter Verwendung der jeweils gleichen Messparameter zu jeder vollen Minute. Die Aufnahmen wurden 30 s nach der Kontrastmittelinjektion begonnen. Von jeder so entstandenen Bildfolge wurde das native T1 Bild (vor Kontrastmittelgabe) subtrahiert. Abschließend erfolgte die Messung einer T2 gewichteten Sequenz sowie einer TIRM-Sequenz. Der Kontrast für freie Flüssigkeit (e. g. perifokale Ödeme im Rahmen von entzündlichen Prozessen) ist in der TIRM-Sequenz höher als im T2 gewichteten Bild, zu Ungunsten der räumlichen Auflösung (Bydder und Young 1985).

Für die klinische Beurteilung der MR-Mammographien wurden neun Bildserien auf einem geeigneten Monitor parallel angesehen. Es handelte sich dabei zum einen um die native T1 Sequenz und T1 gewichtete Serien je zur ersten, zweiten und siebten Minute sowie deren Subtraktionsbilder nach Kontrastmittelgabe. Zum anderen wurden die T2 gewichtete und die TIRM-Sequenz betrachtet.

#### 5.3.4 Grund für die MR-Untersuchung

Allgemein anerkannte, da durch Studien in ihrem Nutzen belegte, Indikationen für die MR-Mammographie sind folgende (Kuhl 2007a):

- Ein unklares, schwer befundbares Röntgenmammogramm (Heywang-Köbrunner et al. 1993)
- bei Tumoren mit unbekanntem Primum (CUP-Syndrom, Cancer of Unknown Primary) wird bei negativem Mammogramm- und Sonographiebefund die MR-Mammographie zur Ausschlussdiagnostik angestrebt (Schorn et al. 1999)
- die Kontrolle der Wirksamkeit der Therapie vor oder nach neoadjuvanter Chemotherapie (Rieber et al. 2002)
- Das Vorhandensein von Brustimplantaten, da hier die Beurteilbarkeit durch die Röntgenmammographie nicht möglich ist
- Die Bewertung (*Staging*) einer bekannten malignen Neoplasie der Mamma (Bedrosian et al. 2003)

Alle Patienten der Kohorte wurden von der Klinik für Frauenheilkunde und Geburtshilfe der FSU Jena zu der MR-Mammographie überwiesen. Tabelle 8 stellt die dabei jeweils angegebenen Gründe für die Überweisung anteilig dar.

### 5.4 Statistische Analyse

#### 5.4.1 Beurteilung der Güte diagnostischer Verfahren

Ein idealer diagnostischer Test sollte mit absoluter Sicherheit zwischen kranken und gesunden Individuen unterscheiden können. Solch ein Test ist real nicht existent. Gründe dafür sind Variationen in der Technik von Untersuchung zu Untersuchung, Limitationen der Technik und die Subjektivität des jeweiligen Untersuchers. Wählt

Tab. 8: GRÜNDE FÜR DIE ÜBERWEISUNG ZU DER MR-MAMMORGRAPHIE

	Anzahl	Anteil in %
unklarer Mammographiebefund	131	17,8
suspekter Mammographiebefund	86	11,7
nur Sonographie suspekt	40	5,4
nur Galaktographie suspekt	1	0,1
Tumornachsorge*	213	28,9
Verlaufskontrolle bei benignem Befund	93	12,6
andere Seite war relevant	154	20,9
nicht eruierbar	18	2,4
Total	736	100

Die Zahlen sind auf die einzelnen untersuchten Mammae bezogen.

\*Unter die Gruppe Tumornachsorge fallen alle Untersuchungen von Patienten, die unter oder nach neoadjuvanter Chemotherapie oder sonstiger abgeschlossener Behandlung einer malignen Neoplasie der Mamma untersucht wurden. Ebenfalls enthalten ist ein CUP-Syndrom. Unter Chemotherapie wurden unabhängig von der Lateralität des Malignoms beide Mammae dieser Gruppe zugeordnet.

man ein quantitativ bestimmbares Merkmal zur Diagnostik, legt ein Schwellenwert dieses Merkmals fest, ab wann ein Individuum als krank gilt. Es wird sich auf diese Art nie vermeiden lassen, einen gewissen Anteil von gesunden Personen als krank, und umgekehrt, einen anderen Anteil von kranken Personen als gesund zu bezeichnen. (Siehe 3.2.2 zu den möglichen Folgen falscher Befunde.)

Tabelle 9 zeigt die diagnostischen Verhältnisse eines binären Tests. Sensitivität, in Worten ausgedrückt, bezeichnet die Wahrscheinlichkeit, dass ein Kranker positiv getestet wird, Spezifität dagegen, dass ein Gesunder ein negatives Testergebnis hat. Der positive prädiktive Wert (PPW) ist die Wahrscheinlichkeit, dass ein positiver Befund tatsächlich maligne ist. Analog dazu bezeichnet der negative prädiktive Wert (NPW) die Wahrscheinlichkeit, dass ein negativer Befund tatsächlich benigne ist. Sowohl positiver als auch negativer prädiktiver Wert sind in hohem Maße von der Prävalenz der entsprechenden Erkrankung abhängig. Die diagnostische Genauigkeit ist der Anteil von richtig positiven und negativen Befunden unter allen Befunden.

Die 95 %-Konfidenzintervalle jeweils für Sensitivität und Spezifität wurden mit folgender Gleichung berechnet, wobei  $g_o$  für die obere Grenze,  $g_u$  für die untere Grenze,

$p$  für den relativen Wert von Sensitivität oder Spezifität von eins sowie  $n$  für die jeweilige Grundgesamtheit steht (Gardner und Altman 1986).

$$g_o, g_u = p \pm 1,96 \times \sqrt{\frac{p \times (1 - p)}{n}}$$

**Tab. 9:** KLASSIFIKATION DES KONZEPTS VON SENSITIVITÄT UND SPEZIFITÄT MITTELS EINER KONTINGENZTAFEL

Testresultat	Diagnose		
	krank	gesund	Total
positiv	$a$	$b$	$a + b$
negativ	$c$	$d$	$c + d$
Total	$a + c$	$b + d$	$a + b + c + d$

$a$ : kranke Individuen die vom Test als krank erkannt werden (richtig positiv),  $b$ : gesunde Individuen die vom Test als krank erkannt werden (falsch positiv),  $c$ : kranke Individuen die vom Test nicht als krank erkannt werden (falsch negativ),  $d$ : gesunde Individuen die vom Test als gesund erkannt werden (richtig negativ), Tabelle nach (Cochrane und Holland 1971).

Sensitivität =  $\frac{a}{a+c} \times 100$ , Spezifität =  $\frac{d}{b+d} \times 100$ , PPW =  $\frac{a}{a+b} \times 100$ , NPW =  $\frac{c}{c+d} \times 100$

Da die Mammographien nach der BIRADS-Klassifikation (siehe Tabelle 6 auf Seite 20) bewertet wurden, muss diese zunächst auf die binären Merkmale *benigne* und *maligne* reduziert werden. BIRADS 0,1,2 und 3 wurden als benigne, 4,5 und 6 als maligne gewertet. Man sollte sich bei dieser Einteilung jedoch bewusst sein, dass im Falle von BIRADS 0 gar keine Aussage über die Mamma getroffen wird (und damit konsequenterweise auch kein Malignomverdacht geäußert wird).

Die Festlegung, ob tatsächlich ein Malignom vorlag oder nicht, wurde nach folgenden logischen Bedingungen getroffen ( $\neg$  Histologie bezeichnet hier das Fehlen einer histologischen Abklärung von suspekten oder unklaren Befunden nach der MR-Untersuchung der Mamma).

tatsächlich benigne:

$(\neg \text{Histologie} \wedge \text{Follow-Up unauffällig}) \vee$

$(\neg \text{Histologie} \wedge \text{gestorben an anderen Gründen als Mamma-Karzinom}) \vee$

(benigne Histologie  $\wedge$  Follow-Up unauffällig)  $\vee$   
 (benigne Histologie  $\wedge$  gestorben an anderen Gründen als Mamma-Karzinom)

tatsächlich maligne:

(maligne Histologie)  $\vee$   
 (benigne Histologie  $\wedge$  Follow-Up suspekt)

Die zeitliche Limitation für falsch negative Befunde wurde in 5.1.4 beschrieben. War in den Mammographiefunden keine BIRADS-Klassifikation explizit angegeben, wurde anhand des schriftlichen Befundes eine Bewertung entweder als benigne oder maligne erfasst. Auf diese Weise als unklar bezeichnete mammographische Befunde wurden als maligne bewertet. Grund hierfür war der Umstand, dass ohne die Möglichkeit der MR-Bildgebung nach diesen Röntgenmammographien die histologische Absicherung des unklaren Befundes wahrscheinlich gewesen wäre.

#### 5.4.2 Kappa-Wert

Der Kappa-Wert ist ein Maß für die Kongruenz der Aussagen zweier unabhängiger Untersucher. Der Wert ist insbesondere bei Nominal- und Ordinalskalen von Nutzen. Bei der BIRADS-Klassifikation handelt es sich um eine ordinale Einteilung. Sie wurde zunächst auf die dichotomen Aussagen *benigne* und *maligne* reduziert, siehe Seite 25. Der Kappa-Wert berechnet sich folgendermaßen:

$$\kappa = \frac{p_0 - p_c}{1 - p_c}$$

$p_0$  ist hier der gemessene Wert der Übereinstimmung der beiden Untersucher und  $p_c$  die zufällig zu erwartende Übereinstimmung. Es wurde vorgeschlagen, den Kappa-Wert folgendermaßen zu bewerten:  $\kappa$  zwischen 0 und 0,2: leichte Übereinstimmung,  $\kappa$  zwischen 0,21 und 0,4: ausreichende Übereinstimmung,  $\kappa$  zwischen 0,41 und 0,6: mittelmäßige Übereinstimmung,  $\kappa$  zwischen 0,61 und 0,8: substantielle Übereinstimmung,  $\kappa$  zwischen 0,81 und 1: (fast) perfekte Übereinstimmung, nach (Landis und Koch 1977).

#### 5.4.3 Chi-Quadrat-Test

Der Chi-Quadrat-Test dient der Überprüfung eines bestehenden Zusammenhanges zwischen zwei Variablen. Er lässt sich auf in Kontingenztafeln darstellbare absolute

Häufigkeiten anwenden. Prinzipiell wird hierbei die Negierung der Nullhypothese ( $H_0$ : Es gibt keinen Unterschied in der Ausprägung der einen Variablen in Abhängigkeit der anderen.) angestrebt. Dazu ermittelt man aus den tatsächlich beobachteten Häufigkeiten die zu erwartenden, wenn kein Zusammenhang bestehen würde. Nun summiert man jeweils die quadrierten Abweichung dieser beiden Werte auf.

$$\sum_{\text{alle Zellen}} \frac{(\text{tatsächliche Häufigkeit} - \text{erwartete Häufigkeit})^2}{\text{erwartete Häufigkeit}}$$

Der Abgleich des so ermittelten Wertes mit feststehenden Tabellenwerten (Bland 2000) lässt Schlüsse darüber zu, wie wahrscheinlich ein Zutreffen der Nullhypothese ist.

#### 5.4.4 Mann-Whitney U Test

Es handelt sich hierbei um einen nicht-parametrischen Test, d.h. er setzt bei den zu testenden Variablen keine Normalverteilung voraus. Der Mann-Whitney U Test ist ein Rangfolgetest, die zu testenden Variablen werden also nach ihren Rängen sortiert. Der Test prüft nun, ob die Rangfolge zufällig zustande gekommen ist (und die beiden Variablen aus der gleichen Grundgesamtheit stammen) oder aber eine echte Differenz der jeweiligen Population besteht.

#### 5.4.5 Konfidenzintervalle für Korrelationskoeffizienten

Aus den ermittelten Korrelationskoeffizienten wird zu der Bestimmung der Konfidenzintervalle mittels *Fisher's z*-Transformation ein  $z$ -Wert berechnet. Grund hierfür ist die generelle Annahme, dass die aus zwei Variablen ermittelten Korrelationskoeffizienten nicht normalverteilt sind (Bland 2000).

$$z = \frac{1}{2} \log_e \left( \frac{1+r}{1-r} \right)$$

Für diesen  $z$ -Wert lässt sich folgendermaßen ein 95 %-Konfidenzintervall ermitteln:

$$g_o, g_u = z \pm 1,96 \times \sqrt{\frac{1}{n-3}}$$

Die so errechneten Werte müssen nun wieder rücktransformiert werden, es wird hierzu die folgende Funktion verwendet:

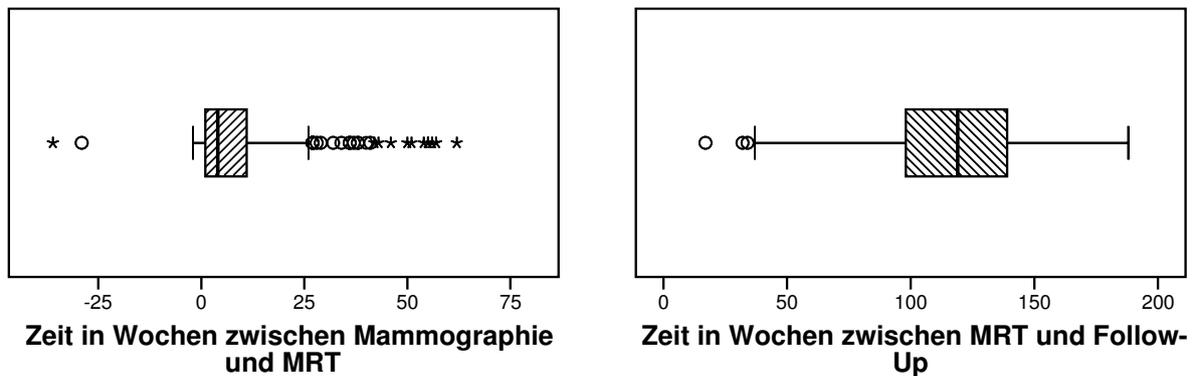
$$r = \frac{\exp(2z) - 1}{\exp(2z) + 1}$$

## 6 Ergebnisse

### 6.1 Ablauf der Studie

#### 6.1.1 Zeitlicher Ablauf

Der mediane zeitliche Abstand zwischen der initialen Röntgenmammographie und der MR-Mammographie betrug vier Wochen, beziehungsweise zwischen MR-Mammographie und Follow-Up-Untersuchung 119 Wochen, siehe Abbildung 9. Damit wurde der angestrebte Zeitabstand von 24 Monaten zwischen der MRM und dem Follow-Up eingehalten. Für die Berechnung des ersten Zeitabstandes wurden 14 Röntgenmammographien ausgeschlossen. Bei diesen handelte es sich um Untersuchungen mit benignem Befund, die Anlass für die mehr als ein Jahr später folgende MR-Untersuchung im Sinne einer Verlaufskontrolle waren.



**Abb. 9:** Darstellung des zeitlichen Abstandes zwischen den jeweiligen Untersuchungen. Der mediane Abstand zwischen Röntgenmammographie und MR-Mammographie betrug vier Wochen, zwischen MR-Mammographie und Follow-Up 119 Wochen.

#### 6.1.2 Anteil ermittelter Röntgenmammographien und Follow-Up-Untersuchungen

Es wurden 19,6% der vorhandenen Follow-Ups sowie 35,3% der initialen, konventionellen Röntgenmammographien von externen Quellen im Rahmen der Studie neu bezogen. Bei 25,4% der Patienten wurde keine initiale Röntgenmammographie angefertigt oder aber deren Befund ließ sich nicht ermitteln. Im Rahmen der Studie wurden von 736 Mammæ MR-Befunde erhoben, zu 676 existierten Follow-Up-Untersuchungen. Analog existierten 549 initiale Röntgenmammographien, von denen 503

über ein Follow-Up verfügten.

In der vorliegenden Arbeit beträgt das Loss-To-Follow-Up 8,4%. In den Tabellen des Ergebnisteils sind fehlende Daten stets mit aufgeführt, um dem Eindruck der statistischen Vollständigkeit vorzubeugen. Die Daten wurden als MNAR (Missing-Not-At-Random (Kristman et al. 2004)) Daten gewertet. Anders ausgedrückt wurde davon ausgegangen, dass die 8,4% der Patienten, die nicht über ein Follow-Up verfügten, sich im Ausgang ihres Follow-Ups von der restlichen Kohorte unterschieden, insbesondere dass prozentual mehr Patienten verstarben oder an einer malignen Neoplasie erkrankten.

## 6.2 Eigenschaften der diagnostizierten Malignome

In der Studienpopulation betrug der Anteil diagnostizierter maligner Neoplasien 37,7%, d.h. bei 342 Patienten wurden 129 maligne Neoplasien histopathologisch gesichert. Die 129 Malignome betrafen 116 Patienten. Da bei der Berechnung von Sensitivität und Spezifität der Mammographie stets von der Anzahl der Mammogramme ausgegangen wird (Smith-Bindman et al. 2005), soll auch hier die Anzahl der untersuchten Mammæ zugrunde gelegt werden. Die 129 malignen Neoplasien verteilten sich damit auf 736 Mammæ, was einem Anteil von 17,5% entspricht. Tabelle 10 gibt eine Übersicht über die Art der diagnostizierten malignen Tumoren.

## 6.3 Diagnostische Genauigkeit der Röntgenmammographie

Bei 25,4% (187 von 736) der mittels MRM untersuchten Mammæ existierte kein Befund der initialen Röntgenmammographie, d. h. es existierten insgesamt 549 initiale Röntgenbefunde. Entweder erhielten diese Patienten die MR-Untersuchung zur Tumornachsorge ohne vorherige Mammographie (11,3%), kamen zur Verlaufskontrolle eines benignen Befundes in die MRM (5,8%), die Mammographie existierte nicht, weil nur die kontralaterale Seite mammographisch von Interesse war (6,5%) oder aber es gab einen anderweitigen Grund für die MR-Überweisung (1,8%).

Tabelle 11 stellt die diagnostischen Ergebnisse der Röntgenmammographie in Form einer Kontingenztafel dar. Aus dieser Aufstellung folgt eine Sensitivität von 64,8% (95%-Konfidenzintervall: 55,7% - 73,9%), eine Spezifität von 93,0% (95%-Konfidenzintervall: 90,5% - 95,5%), ein positiver prädiktiver Wert von 70,8% (95%-Konfidenzintervall: 61,7% - 79,9%) sowie ein negativer prädiktiver Wert von 90,9% (95%-Konfidenzintervall: 88,1% - 93,7%). Die diagnostische Genauigkeit beträgt 87,1% (68 +

**Tab. 10:** VERTEILUNG DER DIAGNOSTIZIERTEN MALIGNOME IN DER STUDIENPOPULATION, DIE KLASSIFIZIERUNG ERFOLGTE HISTOPATHOLOGISCH

Malignom-Typ	Anzahl	Anteil in %	
IDC	85	65,9	} 85,2% invasive Karzinome
ILC	19	14,7	
Mischtyp IDC und ILC	1	0,8	
andere*	15	11,6	
DCIS	9	7,0	
Total	129	100	

IDC: invasiv-duktales Karzinom, ILC: invasiv-lobuläres Karzinom DCIS: duktales carcinoma-in-situ

\*andere steht für: 3 Lymphome, jeweils ein Morbus Paget, invasives Stromasarkom, tubuläres Adenokarzinom, medulläres Karzinom, inflammatorisches Karzinom und muzinöses Karzinom. Sechs Malignome waren nicht typisiert.

370 von 503, 95 %-Konfidenzintervall: 84,2 % - 90,0 %).

Ein entscheidender Punkt bei der Bewertung der diagnostischen Güte ist auch die Vergabe der BIRADS-Klasse 3. Es besteht hier die Wahrscheinlichkeit für ein Malignom von  $\leq 2\%$ . Um zu prüfen, ob es diagnostisch sinnvoller gewesen wäre, BIRADS 3 als maligne zu werten, wurde eine ROC Analyse durchgeführt, siehe Abbildung 10. Tabelle 12 zeigt die zugehörigen statistischen Kennzahlen. Von den 549 mittels BIRADS-Klassen befundeten Röntgenmammographien erhielten 74 die BIRADS-Klasse 3 (13,5 %). Hierunter waren 51 tatsächlich benigne und 16 tatsächlich maligne Befunde, in sieben Fällen war keine Follow-Up-Untersuchung vorhanden. Damit betrug der Anteil an Malignomen unter den Fällen der BIRADS-Klasse 3 23,9 % (16 von 67).

### 6.3.1 Problematik der Parenchymdichte

Tabelle 13 zeigt die Vergabe der BIRADS-Klassen (siehe Tabelle 6 auf Seite 20) nach ACR-Typ in der Studienpopulation. Ein Chi-Quadrat-Test zeigte eine Abhängigkeit der diagnostischen Einschätzung von der jeweiligen Parenchymdichte ( $P < 0,001$ ). So ist bei sehr dichtem Drüsengewebe die Beurteilbarkeit der Röntgenmammographie im Vergleich zu lockerem Gewebe weitaus öfter aufgehoben. Tabelle 14 zeigt die Sensitivität und Spezifität der Röntgenmammographie nach ACR-Subgruppen auf-

**Tab. 11:** KONTINGENZTAFEL FÜR DIE DIAGNOSTISCHE GENAUIGKEIT DER RÖNTGENMAMMOGRAPHIE

Mammographiebefund	tatsächliches Malignom*		Total
	ja	nein	
maligne	68	28	96
benigne	37	370	407
Total	105	398	503**

\*Ob ein Malignom vorlag oder nicht, richtete sich nach dem histopathologischen Befund.

\*\*Die Grundgesamtheit von 503 (Differenz von 46 zu allen 549 vorhandenen Mammographien) kommt durch das Fehlen der Follow-Up-Untersuchung bei 8,4% der Patienten zustande. Für das Vorgehen, anhand dessen in tatsächlich maligne und benigne Läsionen unterschieden wurde, siehe Seite 25.

**Tab. 12:** STATISTISCHE CHARAKTERISTIKA BEI DER KLASSIFIKATION MALIGNER BEFUNDE BEI UNTERSCHIEDLICHER BEWERTUNG VON BIRADS-KLASSE 3 ( $n = 67$ )

	Sensitivität in %	Spezifität in %	Summe von Sensitivität und Spezifität - 1
BIRADS >3 maligne	64,8 (55,7-73,9)	93,0 (90,5-95,5)	0,578*
BIRADS ≥3 maligne	80,0 (72,3-87,7)	80,2 (76,3-84,1)	0,602*

In Klammern hinter Sensitivität und Spezifität sind jeweils die 95%-Konfidenzintervalle angegeben. Die Summe der beiden Kennzahlen bezieht sich auf die jeweiligen relativen Werte von eins.

\* kein signifikanter Unterschied, Test nach (Youden 1950).

geschlüsselt. Für eine Beschreibung der Drüsenparenchyms nach dem ACR siehe Tabelle 3 auf Seite 11.

Untersucht man ausschließlich die Mammae, denen die BIRADS-Klasse 0 gegeben wurde ( $n = 69$ , vergleiche die Anzahl der Fälle aus Tabelle 13. Es wurden vier weitere Fälle aufgenommen, die zwar über eine BIRADS-Klassifikation, jedoch nicht über eine ACR-Klasse verfügten.), so wurde in fünf Fällen zeitnah nach der Röntgenmammographie ein Malignom detektiert, entweder also in einer histologischen Probe, der MRM oder innerhalb eines Jahres nach der MR-Mammographie. Das bedeutet, dass in alleine 7,2% der BIRADS 0 Fälle ein vorhandenes Malignom in der Röntgenmammographie nicht detektiert wurde.

**Tab. 13:** BEWERTUNG DER RÖNTGENMAMMOGRAPHIE NACH ACR-KLASSE

ACR-Klasse	BIRDAS-Klasse			Total
	0	1-3	4-6	
1	4	22	14	40
2	6	105	38	149
3	15	83	56	154
4	40	30	23	93
Total	65	240	131	436*

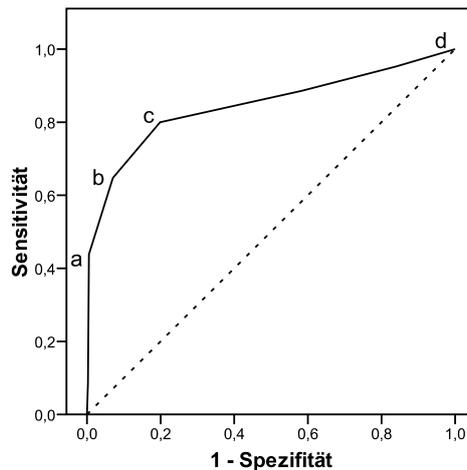
\*Insgesamt war die Dichte des Drüsenparenchyms von 436 Mammae röntgenologisch bestimmt worden, d. h. von allen röntgenologisch befundeten 549 Mammae wurden die verbleibenden 113 zwar nach BI-RADS Kriterien auf Malignität hin geprüft, jedoch ohne Angabe der Dichte des Drüsenparenchyms.

Die BIRADS-Klassen 1-3 werden als benigne, 4-6 als maligne angenommen.

**Tab. 14:** SENSITIVITÄT UND SPEZIFITÄT DER RÖNTGENMAMMOGRAPHIE AUFGESCHLÜSSELT NACH ACR-KLASSEN

ACR-Klasse	Sensitivität in %	Spezifität in %	Mammae	Malignome
1	87,5	93,3	38	8
2	52,9	92,7	141	17
3	70,8	95,1	136	34
4	33,3	98,7	89	12
Total			404*	71

\*Die Differenz von 32 zu allen 436 Mammae (siehe Tabelle 13), aus deren Röntgenbefund sich eine ACR-Klasse eruieren ließ, kommt hier durch das Fehlen von Follow-Up-Untersuchungen zustande. Diese sind notwendig, um mit Sicherheit Sensitivität und Spezifität bestimmen zu können.



**Abb. 10:** Dargestellt ist die ROC (Receiver Operating Characteristic) Kurve. Die unterschiedlichen Werte für Sensitivität und Spezifität kommen durch die Verschiebung der Grenze zustande, ab welcher BIRADS-Klasse ein Befund als maligne bezeichnet wurde. Tabelle 12 zeigt diese unterschiedlichen statistischen Kennzahlen. *a*: BIRADS > 4 maligne, *b*: BIRADS > 3 maligne, *c*: BIRADS  $\geq$  3 maligne, *d*: BIRADS  $\geq$  2 maligne

## 6.4 Diagnostische Genauigkeit der MR-Mammographie

Analog zu der konventionellen Röntgenmammographie lässt sich die diagnostische Genauigkeit der MR-Mammographie mittels einer Kontingenztabelle berechnen, siehe Tabelle 15.

**Tab. 15:** KONTINGENZTAFEL FÜR DIE DIAGNOSTISCHE GENAUIGKEIT DER MR-MAMMOGRAPHIE

MRM-Befund	tatsächliches Malignom		Total
	ja	nein	
maligne	125	33	158
benigne	4	514	518
Total	129	547	676*

\*Die Differenz zu allen 736 MR-mammographisch untersuchten Mammae kommt durch das Fehlen von Follow-Up-Untersuchungen in 60 Fällen zustande.

Die logischen Bedingungen zur Ermittlung tatsächlicher Malignome von Seite 25 gelten weiterhin. Es folgt eine Sensitivität von 96,9 % (95 %-Konfidenzintervall: 93,9 %

- 99,9 %), eine Spezifität von 94,0 % (95 %-Konfidenzintervall: 92,0 % - 96,0 %), ein positiver prädiktiver Wert von 79,1 % (95 %-Konfidenzintervall: 72,8 % - 85,4 %), ein negativer prädiktiver Wert von 99,3 % (95 %-Konfidenzintervall: 98,4 % - 100,0 %) sowie eine diagnostische Genauigkeit von 94,5 % (95 %-Konfidenzintervall: 92,8 % - 96,2 %). Um einen eventuellen Einfluss der Dichte des Brustdrüsenparenchyms auf die diagnostische Genauigkeit der MR-Bildgebung darzustellen, wurden in Tabelle 16 jeweils Sensitivität und Spezifität nach ACR-Klassen aufgeschlüsselt dargestellt.

**Tab. 16:** SENSITIVITÄT UND SPEZIFITÄT DER MR-MAMMOGRAPHIE AUFGESCHLÜSSELT NACH ACR-KLASSEN

ACR-Klasse	Sensitivität in %	Spezifität in %	Mammae	Malignome
1	100,0	100,0	38	8
2	94,1	95,2	141	17
3	94,1	93,1	136	34
4	100,0	94,8	89	12
Total			404	71

Untersucht wurden dieselben Mammae, anhand derer die Sensitivität und Spezifität der Röntgenmammographie in Abhängigkeit von der Parenchymdichte dargestellt wurden, siehe Tabelle 14.

## 6.5 Abweichende Befunde in Röntgenmammographie und MRM

Tabelle 17 betrachtet die gleichen 503 Untersuchungen anhand derer die Sensitivität und Spezifität der Röntgenmammographie bestimmt wurde, nun unter dem erweiterten Gesichtspunkt der Bewertung der MR-Untersuchung. Die Röntgenmammographie wertete dabei 5,5 % aller von der MRM als gutartig bezeichneten Befunde als maligne (23 von 376) und hatte in 8 % dieser Fälle recht mit der Aussage. Andererseits wertete die Röntgenmammographie 42,5 % aller von der MRM als bösartig bezeichneten Befunde als gutartig (54 von 127) und hatte in 33,3 % der Fälle recht. Dies bedeutet das die Röntgenmammographie insgesamt 28 falsch positive und 37 falsch negative Befunde in 503 Untersuchungen erbrachte. Die MR-Mammographie ergab hingegen 25 falsch positive und 3 falsch negative Befunde in dem gleichen Untersuchungskollektiv.

**Tab. 17:** DARSTELLUNG DER ERGEBNISSE JEWEILS VON MRM UND RÖNTGENMAMMOGRAPHIE FÜR DIESELBE MAMMA

MRM-Befund			Malignom		Total
			nein	ja	
benigne	BIRADS, konventionelles Röntgen*	benigne	352	1	353
		maligne	21	2	23
	Total		373	3	376**
maligne	BIRADS konventionelles Röntgen*	benigne	18	36	54
		maligne	7	66	73
	Total		25	102	127**

\*die den konventionellen Mammographiebefunden entnommenen BIRADS-Klassifikationen, auf die binären Aussagen *benigne* und *maligne* reduziert, BIRADS 0,1,2 und 3 wurden als gutartig gewertet

\*\*Die Grundgesamtheit der untersuchten Mammae betrug 503 (376 + 127). Es handelt sich um die gleichen Mammae, anhand derer die diagnostische Genauigkeit der Röntgenmammographie bestimmt wurde, siehe Tabelle 11 auf Seite 32.

## 6.6 Malignomentwicklung und falsch negative Befunde

### 6.6.1 Werdegang detektierter Malignome

Insgesamt entwickelten sich in der Studienpopulation nach einer unauffälligen initialen MRT-Untersuchung sieben maligne Neoplasien neu. Davon traten vier bei Patienten auf, die bereits initial zur ipsilateralen Tumornachsorge die MR-Mammographie erhalten hatten und können damit sämtlich als ipsilaterale Rezidive bezeichnet werden. Der mittlere Zeitabstand zwischen der Detektion des Rezidivs und der ursprünglichen MR-Mammographie betrug 25 Monate (Standardabweichung: 7 Monate). Bei den übrigen drei Malignomen handelte es sich um zwei Ersterkrankungen, die 18 beziehungsweise 30 Monate nach der MR-Mammographie auftraten. In einem Fall entwickelte sich ein kontralaterales inflammatorisches Karzinom, 24 Monate nachdem die MRM das Primum initial detektiert hatte.

Tabelle 18 zeigt eine Übersicht über die Patienten, die innerhalb des Studienzeitraumes verstarben oder aber deren Malignom zum Zeitpunkt des Follow-Ups noch vital war. Nicht mit in diese letzte Gruppe fallen die Patienten aus dem vorigen Absatz, bei denen eine maligne Neoplasie erst nach der unauffälligen MRT-Untersuchung

diagnostiziert wurde.

Insgesamt traten 129 intramammäre Malignome in der Studienpopulation auf, siehe Tabelle 10. Aus den in Tabelle 18 dargestellten Zahlen wird unter dieser Kenntnis deutlich, dass der Großteil der Erkrankten dank erfolgter Therapie zum Zeitpunkt der Follow-Up-Untersuchung tumorfrei war. Von den 129 Erkrankten verstarben 4,7% an Mamma-Karzinomen, 8,5% der Malignome erwiesen sich mindestens teilweise als therapierefraktär.

**Tab. 18:** ÜBERSICHT DER VERSTORBENEN ODER ZUM ZEITPUNKT DES FOLLOW-UPS NOCH KRANKEN PATIENTEN

342 Patienten*	
Verstorben an Malignom der Mamma	6
Malignom zum Zeitpunkt des Follow-Ups noch vital	11
Verstorben anderweitig	4
Verstorben total	10

\*von diesen verfügten 32 Patienten nicht über eine Follow-Up-Untersuchung

### 6.6.2 Falsch negative MRM-Befunde

Die MR-Mammographie wertete vier Befunde als benigne, die sich in der weiteren Abklärung als maligne herausstellten, d.h. es handelte sich um falsch negative Befunde. Es waren dies zwei DCIS-Herde mit Größen von 1,8 mm und 2,0 mm sowie ein 2 mm messendes ILC. Neben diesen wurde in einem Fall sechs Monate nach der unauffälligen MRM ein Morbus Paget (invasiv wachsendes Karzinom der intramammilären Drüsenausführungsgänge) diagnostiziert. Zur Definition der falsch negativen Untersuchung siehe 5.1.4.

## 6.7 Die Messung des Tumordurchmessers

Tabelle 19 zeigt die Korrelationskoeffizienten für die Bestimmung der Durchmesser von Malignomen. Als zutreffende Referenz wurde der von dem Institut für Pathologie der Universitätsklinik Jena angegebene Durchmesser der jeweiligen Neoplasie

angenommen. Es wurden postoperativ zu 82 Malignomen Größenangaben aufgenommen, die mittlere Größe betrug 19,5 mm (Minimum: 2 mm, Maximum: 75 mm). Alle ermittelten Korrelationskoeffizienten zeigen mit  $P < 0,01$  eine signifikante Korrelation zwischen der Größenbestimmung der Neoplasie in der Bildgebung und der postoperative Messung, Signifikanztest nach (Bland 2000). Ein Unterschied zwischen den einzelnen Bildgebungsmodalitäten konnte anhand der 95 %-Konfidenzintervalle nicht festgestellt werden.

**Tab. 19:** KORRELATION DER ANGABE DES TUMORDURCHMESSERS IN DER BILDGEBUNG MIT DER GRÖSSENANGABE DER LÄSION, GEMESSEN NACH DER OPERATIVEN ENTFERNUNG

Bildgebung	Korrelationskoeffizient	Anzahl Neoplasien
Korrelation von allen jeweils gemessenen Neoplasien		
Röntgenmammographie	0,755 (0,56-0,88)	33
MR-Mammographie	0,738 (0,62-0,83)	76
Reevaluierte MR-Mammographie	0,770 (0,67-0,85)	78
Korrelation nur von Neoplasien, zu denen sowohl Röntgenmammographie als auch MR-Mammographie eine Aussage machten		
Röntgenmammographie	0,755 (0,56-0,88)	33
MR-Mammographie	0,898 (0,80-0,95)	33
Reevaluierte MR-Mammographie	0,865 (0,75-0,94)	33

In Klammern sind jeweils die 95 %-Konfidenzintervalle angegeben, zu deren Berechnung siehe 5.4.5.

Es bestand kein signifikanter Unterschied in den Größen der Malignome, die mittels Röntgenmammographie detektiert werden konnten ( $n = 33$ , mittlerer Durchmesser: 19,9 mm) und denen, zu denen ausschließlich die MR-Mammographie einen Durchmesser angab ( $n = 44$ , mittlerer Durchmesser: 20,3 mm). Getestet wurde mit einem Mann-Whitney U Test, es resultierte ein P von 0,68 für den Vergleich der Angaben des Insituts für Pathologie und ein P von 0,32 für den Vergleich der Angaben der reevaluierten MR-Mammographie.

### 6.7.1 Angaben zu Multifokalität und Multizentrität

Die Tabellen 20 bis 22 zeigen die Übereinstimmung der Aussagen von Bildgebung und Pathologie bezüglich des Vorliegens eines monofokalen, multifokalen oder multizentrischen Wachstums. Dargestellt sind hier nur Fälle, in denen sowohl die jeweilige Bildgebungsmodalität als auch das Institut für Pathologie eine Aussage über die Tumorausdehnung machten. Zur Erklärung der Begriffe Monofokalität, Multifokalität und Multizentrität siehe Seite 9.

**Tab. 20:** VERGLEICH DER ANGABEN VON MONOFOKALITÄT, MULTIFOKALITÄT UND MULTIZENTRIZITÄT IN DER RÖNTGENMAMMOGRAPHIE MIT DEN HISTOPATHOLOGISCHEN BEFUNDEN

Mammographie	Pathologie			Total
	monofokal	multifokal	multizentrisch	
monofokal	30	6	1	37
multifokal	4	5	0	9
multizentrisch	0	0	2	2
Total	34	11	3	48

**Tab. 21:** VERGLEICH DER ANGABEN VON MONOFOKALITÄT, MULTIFOKALITÄT UND MULTIZENTRIZITÄT IN DER MR-MAMMOGRAPHIE MIT DEN HISTOPATHOLOGISCHEN BEFUNDEN

MR-Mammographie	Pathologie			Total
	monofokal	multifokal	multizentrisch	
monofokal	42	4	0	46
multifokal	16	11	1	28
multizentrisch	11	14	10	35
Total	69	29	11	109

## 6.8 Die Reevaluation der MR-Mammographien

### 6.8.1 Diagnostische Genauigkeit der Reevaluation

Die Bestimmung von Sensitivität und Spezifität ist mittels einer Kontingenztafel möglich. Beachtet werden muss, dass die Grundgesamtheit der Patienten auf die Fälle

**Tab. 22:** VERGLEICH DER ANGABEN VON MONOFOKALITÄT, MULTIFOKALITÄT UND MULTIZENTRIZITÄT IN DER MR-MAMMOGRAPHIE MIT DEN HISTOPATHOLOGISCHEN BEFUNDEN IN EINEM SELEKTIERTEM PATIENTENKOLLEKTIV (ALLE FÄLLE, ZU DENEN AUCH DIE KONVENTIONELLE RÖNTGENMAMMOGRAPHIE EINE AUSSAGE MACHTE)

MR-Mammographie	Pathologie			Total
	monofokal	multifokal	multizentrisch	
monofokal	22	2	0	24
multifokal	7	7	0	14
multizentrisch	5	1	3	9
Total	34	10	3	47*

\*Es handelt sich hierbei um die gleichen Patienten, deren Daten auch Tabelle 20 zugrunde liegen. In einem Fall hat die MRM keine Aussage über das Ausbreitungsmuster gemacht.

beschränkt ist, die in die Auswahl der Nachevaluation fielen. Dies waren alle Befunde, die histopathologisch vor oder nach der MR-Mammographie abgesichert wurden und vier falsch negative MRM-Befunde. Hinzu kamen alle Mammæ, in denen sich im Zeitraum zwischen MR-Untersuchung und dem Follow-Up ein Malignom neu entwickelte. Insgesamt wurden 216 Mammæ reevaluiert. In den folgenden Kontingenztafeln wurde fünf reevaluierte Untersuchungen ausgeschlossen – eine wegen der Fragestellung nach einem extramammären Tumor, zwei weil der Herd operativ zwischen Röntgenmammographie und MR-Mammographie entfernt wurde und zwei eines Patienten, der zweimal in kurzen Abständen eine MR-Untersuchung erhielt. Die Tabellen 23 bis 25 zeigen die Verteilung jeweils richtiger und falscher Befunde auf das so vorselektierte Patientenkollektiv. Zu 198 der reevaluierten MR-Mammographien lagen korrespondierende Röntgenmammographien vor.

Tabelle 26 zeigt eine Übersicht über die relevanten statistischen Kennzahlen inklusive Konfidenzintervalle für die reevaluierten MR-Mammographien, die ursprünglichen MR-Mammographien sowie die Röntgenmammographien. Es wurde die BIRADS-Bewertung des erfahrenen Befunders für die Reevaluation der MR-Mammographien verwendet. Der Vergleich der beiden Befunder, bei Wertung von BIRADS 1,2 und 3 als benignen Befund und BIRADS 4,5 und 6 als malignen Befund, erbrachte eine fast komplette Übereinstimmung ihrer Einschätzungen (sechs von 211 Mammæ wurden unterschiedlich bewertet,  $\kappa = 0,938$ ).

**Tab. 23:** KONTINGENZTAFEL FÜR DIE DIAGNOSTISCHE GENAUIGKEIT DER GEBLINDET REEVALUIERTEN MR-MAMMOGRAPHIEN

MRM-Befund	tatsächliches Malignom		Total
	ja	nein	
maligne	119	11	130
benigne	6	62	68
Total	125	73	198*

\*Die Differenz zu den eingeschlossenen 211 Untersuchungen kommt durch das Fehlen von Follow-Up-Untersuchungen in 13 Fällen zustande. Von allen 216 reevaluierten MR-Mammographien wurden fünf für diese Analyse ausgeschlossen: eine wegen der Fragestellung nach einem extramammären Tumor, zwei weil der Herd operativ zwischen Röntgenmammographie und MR-Mammographie entfernt wurde und zwei eines Patienten, der zweimal in kurzen Abständen eine MR-Untersuchung erhielt.

**Tab. 24:** KONTINGENZTAFEL FÜR DIE DIAGNOSTISCHE GENAUIGKEIT DER ERSTBEFUNDETEN MR-MAMMOGRAPHIE FÜR DAS NACH TABELLE 23 SELEKTIERTE PATIENTENKOLLEKTIV

MRM-Befund	tatsächliches Malignom		Total
	ja	nein	
maligne	121	33	154
benigne	4	40	44
Total	125	73	198*

\*Die Differenz zu den eingeschlossenen 211 Untersuchungen kommt durch das Fehlen von Follow-Up-Untersuchungen in 13 Fällen zustande.

### 6.8.2 Beschreibung der reevaluierten Herdbefunde

In den 211 reevaluierten Mammae (zum Zustandekommen der Grundgesamtheit siehe 6.8.1) fanden sich 138 mass- und 38 non-mass-Läsionen, zur Unterscheidung dieser beiden Formen siehe 5.2.2. In den übrigen 35 untersuchten Mammae beschrieb die Reevaluation keinen Herd und vergab auf diese die BIRADS-Klasse 1. Ein entscheidendes diagnostisches Kriterium für die Unterscheidung von malignen und benignen Herdbefunden ist die Kinetik der Kontrastmittelanreicherung des Herdes. Tabelle 27 zeigt die Verteilung der verschiedenen Kontrastmittelkinetiken auf mass- und non-mass-Läsionen, siehe ebenfalls Abbildung 11 auf Seite 44.

**Tab. 25:** KONTINGENZTAFEL FÜR DIE DIAGNOSTISCHE GENAUIGKEIT DER RÖNTGENMAMMOGRAPHIE FÜR DAS NACH TABELLE 23 SELEKTIERTE PATIENTENKOLLEKTIV

Mammographie-Befund	tatsächliches Malignom		Total
	ja	nein	
maligne	67	12	79
benigne	35	46	81
Total	102	58	160*

\*Die Differenz zu der Anzahl der eingeschlossenen Untersuchungen in Tabelle 23 kommt durch das Fehlen von Mammographie-Aufnahmen in diesen Fällen zustande (38 von 198).

### 6.8.3 Falsch negative MR-Befunde

Insgesamt werteten die Befunder der ursprünglichen MR-Mammographie 37 Untersuchungen falsch, vier falsch negativ und 33 falsch positiv, siehe Tabelle 24. Die vier falsch negativen Befunde entsprechen den in 6.6.2 dargestellten. In der Reevaluation war es den beiden verblindeten Untersuchern ebenfalls nicht möglich, die ILC- und die beiden DCIS-Läsionen (alle kleiner als drei Millimeter) zu detektieren. Zusätzlich wurden ein multifokales, je Herd höchstens vier Millimeter im Durchmesser messendes DCIS des Malignitätsgrades eins, ein 20 mm messendes Stromasarkom sowie ein invasiv duktales Karzinom nach abgeschlossener Chemotherapie nicht detektiert. Den Fall des Morbus Paget erkannten die reevaluierenden Befunder hingegen als maligne.

### 6.8.4 Gründe für falsch positive Befunde in der MRM

Bei den 33 falsch positiven Befunden der MR-Mammographie handelte es sich, den Ergebnissen der Reevaluation folgend, um 17 mass- und 14 non-mass-Läsionen. In zwei Fällen ließ sich kein Herd vom Drüsenparenchym abgrenzen. Dies bedeutet, dass das Verhältnis von mass- zu non-mass Herden unter den falsch positiven Fällen 17 zu 14 (1,2), unter allen Herden der Studienpopulation aber 138 zu 38 (3,6) betrug. Die Tabellen 28 und 29 zeigen die Ergebnisse der Diagnostik von originärer und reevaluierter MRM für alle non-mass-Läsionen. Es wird deutlich, dass hier im Falle einer malignen Diagnose der ursprünglichen MR-Mammographie diese keine signifikante Hilfe zur Entscheidungsfindung war.

**Tab. 26:** STATISTISCHE KENNZAHLEN FÜR DIE REEVALUIERTEN MR-MAMMOGRAPHIEN, DIE URSPRÜNGLICHEN MR-MAMMOGRAPHIEN SOWIE DIE RÖNTGENMAMMOGRAPHIEN

	geblindet reevaluierte MR-Mammographie	ursprüngliche MR-Mammographie	Röntgen- mammographie
Sensitivität in %	95,2 (91,5 - 98,9)	96,8 (93,7 - 99,9)	65,7 (56,5 - 74,9)
Spezifität in %	84,9 (76,6 - 93,1)	54,8 (43,4 - 66,2)	79,3 (68,9 - 89,7)
PPW in %	91,5 (86,7 - 96,3)	78,6 (72,1 - 85,1)	84,8 (76,9 - 92,7)
NPW in %	91,2 (84,5 - 97,9)	90,9 (82,4 - 99,4)	56,8 (46,0 - 67,8)
Diagnostische Genauigkeit in %	91,4 (87,5 - 95,3)	81,3 (75,9 - 86,7)	70,6 (63,5 - 77,7)

Die Werte beziehen sich auf die in den Tabellen 23 bis 25 verwendeten Untersuchungen. In Klammern sind jeweils die 95 %-Konfidenzintervalle angegeben. PPW: positiver prädiktiver Wert, NPW: negativer prädiktiver Wert

Um nun anhand der BIRADS-Kriterien (siehe Tabelle 5 auf Seite 17) mögliche Deskriptoren festzustellen, die sich in falsch positiven Befunden überproportional häufig finden, ist der direkte Vergleich mit den zutreffenden Befunden notwendig.

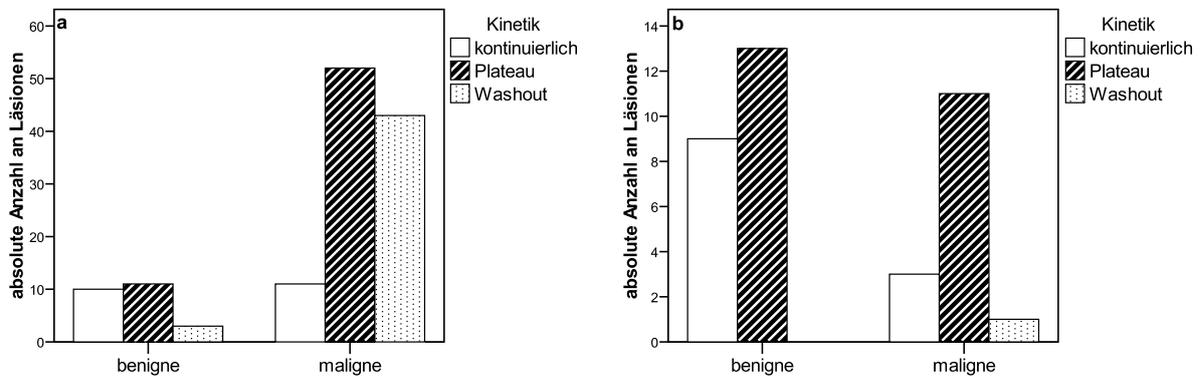
Der gemessene Durchmesser in der Bildgebung hatte keinen Einfluss auf eine mögliche Fehldiagnose im Falle von non-mass-Tumoren. Verglichen wurden zu diesem Zweck jeweils die Durchmesser der falsch positiven und der tatsächlich malignen Läsionen, die in der MR-Reevaluation gemessen wurden. Der Test auf einen signifikanten Unterschied erfolgte mit einem Mann-Whitney U Test. Für non-mass-Tumoren betrug P 0,306 (14 falsch positive und 14 tatsächlich maligne Läsionen wurden eingeschlossen). Im Falle der Mass-Läsionen unterschieden sich die Größen von falsch positiven und tatsächlich malignen Läsionen signifikant. Ein Mann-Whitney U Test zeigte für diesen Unterschied ein P von 0,012 (17 falsch positive und 106 tatsächlich maligne Läsionen wurden eingeschlossen). Die mittlere Größe der falsch positiven Befunde betrug 16,9 mm, die der Malignome hingegen 23,6 mm.

Die Tabellen 30, 31 und 32 zeigen die Verteilung der BIRADS-Deskriptoren auf die tatsächlich malignen sowie falsch positiv und richtig negativ bewerteten non-masses. Die Tabellen 33 und 34 geben diese für non-mass-Läsionen an.

**Tab. 27:** VERTEILUNG DER VERSCHIEDENEN KONTRASTMITTEL-KINETIKEN AUF MALIGNEN UND BENIGNEN MASS- UND NON-MASS-LÄSIONEN

Kinetik	mass		non-mass		Total
	benigne	maligne	benigne	maligne	
kontinuierlich	10	11	9	3	33
Plateau	11	52	13	11	87
Washout	3	43	0	1	47
Total	24	106	22	15	167*

\*Es sind hier 130 mass- und 37 non-mass-Läsionen enthalten. Die jeweilige Differenz zu allen reevaluierten Herden kommt durch das Fehlen der Follow-Up-Untersuchung in den entsprechenden Fällen zustande. Gutartige Herdbefunde wurden nicht in die Übersicht aufgenommen.



**Abb. 11:** In Abbildung *a* ist die Verteilung der verschiedenen Anreicherungsverhalten in mass-Läsionen zu sehen, Abbildung *b* auf der rechten Seite zeigt die Verhältnisse bei non-mass-Läsionen.

**Tab. 28:** KONTINGENZTAFEL FÜR DIE DIAGNOSTISCHE GENAUIGKEIT DER URSPRÜNGLICHEN MR-MAMMOGRAPHIE FÜR SÄMTLICHE NON-MASS-LÄSIONEN (SELEKTIONIERTES PATIENTENKOLLEKTIV)

MRT-Befund	tatsächliches Malignom		Total
	ja	nein	
maligne	14	14	28
benigne	0	9	9
Total	14	23	37*

\*Es handelt sich um sämtliche non-mass-Läsionen, die auch Tabelle 27 auf Seite 44 zugrunde lagen.

**Tab. 29:** KONTINGENZTAFEL FÜR DIE DIAGNOSTISCHE GENAUIGKEIT DER GEBLINDET REEVALUIERTEN MR-MAMMOGRAPHIE FÜR SÄMTLICHE NON-MASS-LÄSIONEN (SELEKTIONIERTES PATIENTENKOLLEKTIV)

MRT-Befund	tatsächliches Malignom		Total
	ja	nein	
maligne	13	6	19
benigne	1	17	18
Total	14	23	37*

\*Es handelt sich um sämtliche non-mass-Läsionen, die auch Tabelle 27 auf Seite 44 zugrunde lagen.

**Tab. 30:** DIE HERDKONFIGURATION MALIGNER SOWIE ALS FALSCH POSITIV UND FALSCH NEGATIV BEWERTETER BEFUNDE (MASS-LÄSIONEN)

Konfiguration	falsch positiv	maligne	richtig negativ
rund	6 (35,3 %)	20 (18,9 %)	1 (14,3 %)
oval	3 (17,6 %)	16 (15,1 %)	2 (28,6 %)
lobuliert	7 (41,2 %)	52 (49,1 %)	4 (57,1 %)
irregulär	1 (5,9 %)	18 (17,0 %)	0 (0,0 %)
Total	17	106	7

Es handelt sich um die gleichen mass-Läsionen, die auch Tabelle 27 auf Seite 44 zugrunde liegen: 106 maligne sowie 24 benigne Läsionen.

**Tab. 31:** DIE HERDBERANDUNG MALIGNER SOWIE ALS FALSCH POSITIV UND FALSCH NEGATIV BEWERTETER BEFUNDE (MASS-LÄSIONEN)

Berandung	falsch positiv	maligne	richtig negativ
glatt	10 (58,8 %)	4 (3,8 %)	6 (85,7 %)
irregulär	6 (35,3 %)	72 (67,9 %)	1 (14,3 %)
spikuliert	1 (5,9 %)	30 (28,3 %)	0 (0,0 %)
Total	17	106	7

Es handelt sich um die gleichen mass-Läsionen, die auch Tabelle 27 auf Seite 44 zugrunde liegen: 106 maligne sowie 24 benigne Läsionen.

**Tab. 32:** DAS ANREICHERUNGSMUSTER MALIGNER SOWIE ALS FALSCH POSITIV UND FALSCH NEGATIV BEWERTETER BEFUNDE (MASS-LÄSIONEN)

Anreicherung	falsch positiv	maligne	richtig negativ
homogen	0 (0,0 %)	1 (0,9 %)	1 (14,3 %)
heterogen	13 (76,5 %)	65 (61,3 %)	4 (57,1 %)
zentripetal	3 (17,6 %)	36 (34,0 %)	0 (0,0 %)
zentrifugal	1 (5,9 %)	0 (0,0 %)	0 (0,0 %)
dunkle Septierung	0 (0,0 %)	1 (0,9 %)	1 (14,3 %)
anreichernde Septierung	0 (0,0 %)	3 (2,8 %)	1 (14,3 %)
Total	17	106	7

Es handelt sich um die gleichen mass-Läsionen, die auch Tabelle 27 auf Seite 44 zugrunde liegen: 106 maligne sowie 24 benigne Läsionen.

**Tab. 33:** DIE RÄUMLICHE VERTEILUNG MALIGNER SOWIE ALS FALSCH POSITIV UND FALSCH NEGATIV BEWERTETER BEFUNDE (NON-MASS-LÄSIONEN)

Verteilung	falsch positiv	maligne	richtig negativ
fokal	6 (42,9 %)	3 (20,0 %)	5 (62,5 %)
linear	0 (0,0 %)	0 (0,0 %)	0 (0,0 %)
duktal	2 (21,4 %)	1 (6,7 %)	2 (25,0 %)
segmental	3 (5,9 %)	8 (53,3 %)	1 (12,5 %)
regional	2 (14,3 %)	2 (13,3 %)	0 (0,0 %)
multipel regional	1 (7,1 %)	1 (6,7 %)	0 (0,0 %)
Total	14	15	8

$n = 37$ , siehe hierfür Tabelle 27 auf Seite 44

**Tab. 34:** DAS ANREICHERUNGSMUSTER MALIGNER SOWIE ALS FALSCH POSITIV UND FALSCH NEGATIV BEWERTETER BEFUNDE (NON-MASS-LÄSIONEN)

Anreicherung	falsch positiv	maligne	richtig negativ
homogen	0 (0,0 %)	0 (0,0 %)	2 (25,0 %)
heterogen	2 (14,3 %)	6 (40,0 %)	1 (12,5 %)
stippled	6 (42,9 %)	1 (6,7 %)	4 (50,0 %)
clumped	4 (28,6 %)	4 (26,7 %)	0 (0,0 %)
retikulär dendritisch	2 (14,3 %)	4 (26,7 %)	1 (12,5 %)
Total	14	15	8

$n = 37$ , siehe hierfür Tabelle 27 auf Seite 44

## 7 Diskussion

### 7.1 Statistische Vorüberlegungen

Zentrales Anliegen der vorliegenden Studie ist die Bestimmung der diagnostischen Genauigkeit der MR-Mammographie und der direkte Vergleich dieser mit der konventionellen Röntgenmammographie. Vor der näheren Betrachtung einzelner Ergebnisse sollen nun zwei Grundüberlegungen hinsichtlich der diagnostischen Genauigkeit nachvollzogen werden.

Erstens sind die ermittelten statistischen Kennzahlen eines Tests (i. e. Sensitivität, Spezifität, positiver prädiktiver sowie negativer prädiktiver Wert) unmittelbar nur auf die jeweilige Grundgesamtheit anwendbar, anhand derer sie ermittelt wurden. So hängt der positive prädiktive Wert eines diagnostischen Tests maßgeblich von der Prävalenz der entsprechenden Krankheit in der untersuchten Grundgesamtheit ab (Altman und Bland 1994). Je seltener die Erkrankung ist, desto zuverlässiger wird die Aussage des Tests im Falle eines benignen Befundes. Jedoch gilt ebenfalls, dass, je seltener die Erkrankung in der untersuchten Kohorte ist, ein maligner Befund stetig an Aussagekraft verliert.

Zweitens ist das Fehlen von Follow-Up-Untersuchungen kritisch zu betrachten. Aus Tabelle 9 auf Seite 25 geht hervor, dass eine zufriedenstellende Spezifität bei einem diagnostischen Test nur durch eine große Anzahl richtig negativer Befunde ermöglicht werden kann (da es immer einen bestimmten Anteil an falsch positiven Befunden geben wird). Die benignen Befunde können, bei fehlender histopathologischer Herdabklärung, einzig durch vorhandene Follow-Up-Untersuchungen verifiziert werden. Für das Fehlen von Follow-Up-Untersuchungen sind für gewöhnlich systematische Gründe verantwortlich, d. h. die Daten fehlen nicht zufällig (Collins et al. 2001). Betrachtet man die Daten als nicht zufällig verteilt fehlend, muss als Grund ein Unterschied in dem klinischen Verlauf der Patienten im Vergleich zur restlichen Studienkohorte angenommen werden (Twisk und de Vente 2002, Bisgard et al. 1994). In der Studienpopulation betrug der Anteil an malignen MR-Befunden 16,7% unter allen Untersuchungen, für die kein Follow-Up eruiert werden konnte. Für das restliche Patientenkollektiv, welches über ein Follow-Up verfügte, betrug dieser Wert 23,3%. Hierbei sind auch die nachgewiesenen 33 falsch positiven Befunde enthalten. Damit sind überproportional viele Patienten mit benignem Befund ohne Follow-Up. Der vermeintlich logische Schluss, diese Kohorte hätte nun tatsächlich weniger maligne Neoplasien entwickelt, ist jedoch unzulässig.

In der vorliegenden Studie konnten zu 8,4 % der initial vorhandenen MR-Untersuchungen keine Follow-Up-Untersuchungen ermittelt werden. Ab einem Loss-To-Follow-Up von 20 % sind die aus Kontingenztafeln abgeleiteten Verhältnisse mit erheblichen Störungen belastet, unter der Prämisse, die Daten seien nicht rein zufällig verteilt fehlend. Bei einem Verlust von 10 % der Follow-Up-Daten schließen die aus den Kontingenztafeln ermittelten 95 %-Konfidenzintervalle noch rund 93 % der wahren Verhältnisse ein (Kristman et al. 2004). Es darf damit angenommen werden, dass die in der vorliegenden Arbeit ermittelten statistischen Kennzahlen repräsentativ für die Studienkohorte sind. Gängige Praxis ist es, schon ein benignes klinisches Follow-Up als exakte Referenz für initial benigne Befunde zu werten. Jedoch sollte man sich darüber im klaren sein, dass auch die Sensitivität beinahe sämtlicher Follow-Up-Methoden (i. e. MRM, Röntgenmammographie, Ultraschall, klinisches Follow-Up) eben nicht 100 % ist. So könnten sich Malignome zwischen der originären MRM und der Follow-Up-Untersuchung entwickelt haben, die in letzterer nicht detektiert wurden.

Die Prävalenz maligner Neoplasien übersteigt in der vorliegenden Studienkohorte (17,5 %) den Wert der Screeningpopulation (1,4 %) um ein Vielfaches. Die Kohorte der vorliegenden Arbeit rekrutiert sich also nicht zufällig aus der Screeningpopulation. In der kaukasischen Screeningpopulation werden relative Anteile von 78,7 % für invasive beziehungsweise 21,3 % für nichtinvasive Neoplasien angegeben (Smigal et al. 2006). Die Studienpopulation hatte, bei Zuordnung von Lymphomen und nicht näher klassifizierten Tumoren in die Gruppe der nicht-invasiven Formen, einen Anteil von 85,2 % invasiver Malignome unter allen malignen Befunden, siehe Tabelle 10 auf Seite 31. Damit besteht hinsichtlich dieses Aspektes eine gute Vergleichbarkeit beider Populationen. Es wird in der folgenden Diskussion immer wieder der Vergleich zu der Screeningpopulation angestellt werden. Grund hierfür ist letztlich, dass anhand dieser herausforderndsten Situation die Genauigkeit der MRM überprüft werden soll.

## 7.2 Einordnung der ermittelten diagnostischen Genauigkeiten

### 7.2.1 Konventionelle Röntgenmammographie

Sensitivität und Spezifität der konventionellen Röntgenmammographie waren und sind Gegenstand zahlreicher Studien. Die Angaben für die Sensitivität bewegen sich zwischen 63 % und 98 % (Robertson 1993, Burhenne et al. 1994, Kerlikowske et al. 1996). Im Falle dichten Brustdrüsenparenchyms wurden Werte zwischen 30 % und

45 % bestimmt (Mandelson et al. 2000, Kolb et al. 2002). Dies zeigt die Abhängigkeit der Sensitivität von der Dichte des Brustdrüsengewebes. In den zitierten Arbeiten rekrutierte sich die Studienpopulation jeweils aus der Screeningpopulation. Die Angaben bezüglich der Spezifität der Röntgenmammographie bewegen sich zwischen 75 % und 97 % (Duijm et al. 1997, Sardanelli et al. 2004). Der positive prädiktive Wert eines Tests hängt, wie in 7.1 dargelegt, von der Prävalenz der entsprechenden Krankheit in der Studienpopulation ab. In den zitierten Publikationen nimmt der positive prädiktive Wert mit steigender Zahl der Studienteilnehmer stetig ab, er bewegt sich zwischen 18 % und 73 %.

Die in der vorliegenden Arbeit ermittelten statistischen Kennzahlen (Daten aus Tabelle 11 auf Seite 32, Sensitivität: 64,8 %, Spezifität: 93,0 %, positiver beziehungsweise negativer prädiktiver Wert: 70,8 %, 90,9 %, diagnostische Genauigkeit: 87,1 %) bestätigen die aus der Literatur bekannten Werte. Im Vergleich zu wirklichen Screeningstudien spricht der hohe positive prädiktive Wert für die hohe Prävalenz maligner Neoplasien in der Studienpopulation. Ein weit besseres Maß für die Sicherheit eines positiven Befundes ist unter diesen Umständen die *likelihood-ratio* (Altman und Bland 1994). Sie ist definiert durch den Quotienten von Sensitivität und  $1 - \text{Spezifität}$ . Im Falle der Röntgenmammographie ergibt sich hierfür ein Wert von 9,25. Je größer das Verhältnis von Sensitivität zu  $1 - \text{Spezifität}$  ist, desto größer ist die diagnostische Sicherheit eines positiven Befundes. Es besteht ein statistisch signifikanter Zusammenhang zwischen der Parenchymdichte und der Vergabe der BIRADS-Klassen. So wurde die Klasse 0 überproportional häufig an sehr dichte Mammae vergeben. Die Abhängigkeit der diagnostischen Genauigkeit von der Dichte des Brustdrüsenparenchyms konnte ebenfalls gezeigt werden, siehe Tabelle 14 auf Seite 33. Dies geht konform mit der vorherigen Aussage, dass dichtes Brustdrüsengewebe die mammographische Beurteilung in einem großen Teil der Fälle massiv erschwert und dadurch vermehrt zu Fehlbefunden führt.

Tabelle 14 auf Seite 33 zeigt eine bessere Beurteilbarkeit von Fällen mit einer Parenchymdichte von ACR 3 als von Fällen mit ACR 2. Dies erscheint zunächst zu der bekannten Literatur widersprüchlich zu sein. Die Anteil von Malignomen unterschied sich in diesen beiden Subgruppen jedoch um den Faktor zwei (34 aus 136 im Vergleich zu 17 aus 141). Die ermittelten statistischen Kennzahlen sind deshalb nur eingeschränkt für generelle Aussagen geeignet. Auch wird der klare Trend einer abnehmenden Sensitivität mit steigender Drüsenparenchymdichte durch die beiden ermittelten Werte nicht weniger offensichtlich.

Die BIRADS-Klasse 3 wurde in der Studienpopulation an 13,5 % aller nach BIRADS-Kriterien bewerteten Befunde vergeben. Andere Studien geben hier Werte zwischen 3 % und 10 % an (Varas et al. 2002, Orel et al. 1999). Weit wichtiger als die Häufigkeit ist jedoch die Aussagekraft des Befundes. Die BIRADS-Klasse 3 für sich genommen weist in den bisher publizierten Studien jeweils einen enorm niedrigen positiven prädiktiven Wert auf, er wird zwischen 0,4 % und 3 % angegeben (Sickles 1991, Sickles 1999, Orel et al. 1999, Varas et al. 2002, Masroor 2008). Dies bedeutet, dass im Allgemeinen eine mit BIRADS 3 klassifizierte intramammäre Läsion sehr wahrscheinlich benigne ist. In der vorliegenden Arbeit betrug der Anteil an Malignomen unter den Fällen der BIRADS-Klasse 3 23,9 % (16 von 67). Eine Ursache hierfür mag der Umstand sein, dass die Aussagen zu Fällen der BIRADS-Klasse 3 in der genannten Literatur sich auf Fälle beziehen, die sich direkt aus der Screeningpopulation rekrutierten. Die Vorselektion der Patienten in der vorliegenden Arbeit (die gesamte Population erhielt eine MR-Untersuchung, für die Gründe siehe Tabelle 8) ist hinreichend, sowohl die höhere Prävalenz von Malignomen unter allen BIRADS 3 Befunden zu erklären als auch die, relativ gesehen, häufigere Vergabe dieser Klassifikation. Andererseits könnte eine unterschiedliche generelle Praxis in der Vergabe solcher Befunde für den hohen Anteil an Malignomen unter den Fällen der BIRADS-Klasse 3 verantwortlich sein.

In Tabelle 12 auf Seite 32 wurden Sensitivität und Spezifität für die beiden unterschiedlichen Bewertungen der BIRADS-Klasse 3 angegeben. Ein bewährtes Konzept zum direkten Vergleich beider Einordnungen ist in Hinblick auf die durchgeführte ROC Analyse die Summe der beiden Werte minus eins (Youden 1950). Obwohl sich die Spezifität beider diagnostischer Vorgehensweisen signifikant unterscheidet (siehe die 95 % Konfidenzintervalle in Tabelle 12), konnte kein Unterschied in den summierten Proportionen aller richtigen Befunde gefunden werden (Youden 1950). Das wiederum bedeutet, dass es in der Studienpopulation statistisch bedeutungslos ist, ob man BIRADS-Klasse 3 als benigne oder als maligne wertet. Daraus ergibt sich die Konsequenz, Patienten mit einem mammographischen Befund der BIRADS-Klasse 3, die eine MR-Überweisung aufgrund dieses Befundes erhalten, nicht als eine zufällig aus allen Patienten mit Befunden der BIRADS-Klasse 3 gezogene Stichprobe zu betrachten.

### 7.2.2 MR-Mammographie

Die MR-Mammographie ist kein allgemein anerkanntes Screeningverfahren für Malignome der Mamma. Sensitivität und Spezifität, ebenfalls Gegenstand zahlreicher Studien, wurden so in weiten Teilen stets an vorselektierten Patientenkollektiven bestimmt, je nach Indikation für die MRT, siehe 5.3.4. Sämtliche Studien, die bisher MR-Mammographie mit konventioneller Röntgenmammographie verglichen, bedienen sich ebenfalls eines derart vorselektierten Patientenkollektivs. Insbesondere wurden genetisch prädisponierte Patientenkollektive untersucht (Kuhl 2007a). Die Autorin argumentiert, dass sich diese Patienten bis auf die Prävalenz der Malignome nicht fundamental biologisch von der Screeningpopulation unterscheiden. Für weitere Ausführungen zu dem direkten Vergleich der MR-Mammographie mit der Röntgenmammographie siehe 7.2.3. Tabelle 35 zeigt eine Aufstellung aktueller Studien von Sensitivität und Spezifität der MR-Mammographie. Immer wieder wird die stark schwankende beziehungsweise niedrige Spezifität der MR-Mammographie als Grund angeführt, diese Methode nicht zum Screening zu verwenden. Dabei wird außer acht gelassen, dass viele der Studien bezüglich der MR-Mammographie ausschließlich die Daten histopathologisch abgesicherter Herdbefunde verwenden. Selektionsbedingt wird auf diese Weise der Anteil richtig negativer Befunde vermindert. Die Spezifität wird somit künstlich gesenkt.

Die vorliegende Studie ist die erste Arbeit, deren Population sich konsequent aus der momentanen klinischen Anwendung der MR-Mammographie rekrutiert. Diese Population ist also anders vorselektiert als jene der Studien in Tabelle 35. Bei dem überwiegenden Teil der publizierten Studien wird die Aussagekraft durch das Vorliegen von mindestens einem der folgenden Kritikpunkte gestört: Erstens mögen die Follow-Up-Untersuchungen nur in einem unzureichenden Ausmaß vorhanden sein, womit die Verifikation eines entsprechenden Anteiles richtig negativer Befunde verhindert wird. Zweitens wird oftmals die berechtigte Forderung nach der Angabe der Konfidenzintervalle statistischer Kennzahlen nicht erfüllt (Gardner und Altman 1986). Diese sind jedoch zwingend erforderlich, um wenigstens innerhalb der gleichen Studie verschiedene diagnostische Modalitäten befriedigend vergleichen zu können. Beides trifft nicht auf die vorliegende Arbeit zu.

Die ermittelten statistischen Kennzahlen sind in 6.4 aufgeführt. (Sensitivität: 96,9 %, Spezifität: 94,0 %, positiver beziehungsweise negativer prädiktiver Wert: 79,1 %, 99,3 %, diagnostische Genauigkeit: 94,5 %). Insgesamt stützen die statistischen Kennzahlen das beschriebene Bild, es ergibt sich eine *likelihood-ratio* von 16,15. Auffällig ist der

**Tab. 35:** ÜBERSICHT ÜBER STUDIEN, DIE SENSITIVITÄT UND SPEZIFITÄT VON RÖNTGENMAMMOGRAPHIE UND MR-MAMMOGRAPHIE VERGlichen. DIE PATIENTENKOLLEKTIVE HATTEN IN JEDEM FALL EIN MITTELGRADIGES BIS HOCHGRADIGES KUMULATIVES LEBENSZEITRISIKO FÜR MALIGNEN BRUSTERKRANKUNGEN. DIE JEWEILIGEN STATISTISCHEN KENNZAHLEN FÜR DIE RÖNTGENMAMMOGRAPHIE SIND IN TABELLE 36 DARGESTELLT.

Autor, Jahr	Patienten	Malignome	MRT				
			Sensitivität in %	Spezifität in %	PPW in %	NPW in %	Loss-to- follow-up in %
Sardanelli, 2007	278	18	95	-	63 (43-79)	-	64,4 pro 2 a
Kuhl, 2005	618	12	83	-	43	-	k.A.
Kuhl, 2005	529	43	91	97	50	-	23,9 pro 2 a
Leach, 2005	649	35	77 (60-90)	81 (80-83)	7,3 (4,9-10)	99 (99-100)	2,6 pro 1 a
Warner, 2004	236	22	85	93	42	99	0 pro 1 a, 44 pro 2 a
Kriege, 2004	1909	45	80	90	32	-	4,6 pro 1 a

In Klammern sind die 95 % Konfidenzintervalle angegeben, wenn sie sich aus den zitierten Arbeiten ermitteln ließen. (Sardanelli et al. 2007, Kuhl et al. 2005b, Kuhl et al. 2005a, Leach et al. 2005, Warner et al. 2004, Kriege et al. 2004), modifizierte Tabelle nach (Kuhl 2007a).

hohe positive prädiktive Wert. Grund hierfür ist wahrscheinlich die weitaus höhere Prävalenz von Malignomen in der Studienpopulation (129 von 342 Patienten, entspricht 37,7 %) als in den in Tabelle 35 aufgeführten Publikationen (1,9 % - 9,3 %). Der Umstand, dass für die Berechnung von Sensitivität und Spezifität in den aufgeführten Arbeiten der Patient als kleinste Untersuchungseinheit angenommen wurde und nicht die einzelne Mamma, senkt die Zahl richtig negativ befundeter Untersuchungen stark ab. In besonderem Maße erscheint diese Praxis nicht einleuchtend, da der direkte Vergleich mit der Röntgenmammographie hergestellt werden soll, deren Screeningstudien sich auf einzelne Mammogramme beziehen. Tabelle 16 auf Seite 35 bestätigt, dass Sensitivität und Spezifität der MR-Mammographie weitgehend unabhängig von der Parenchymdichte der Mamma sind (Sardanelli et al. 2004, Kawashima et al. 2000).

### 7.2.3 Der direkte Vergleich beider Bildgebungsmodalitäten

Tabelle 36 zeigt die zu den Studien der Tabelle 35 korrespondierenden statistischen Kennzahlen der konventionellen Röntgenmammographie. Es existiert keine Studie, in der die Sensitivität der MR-Mammographie kleiner war als die der Röntgenmammographie (Kuhl 2007b).

**Tab. 36:** KORRESPONDIERENDE ERGEBNISSE DER STUDIEN AUS TABELLE 35, NUN AUF DIE RÖNTGEN-MAMMOGRAPHIE BEZOGEN.

Autor, Jahr	Patienten	Malignome	Mammographie			
			Sensitivität in %	Spezifität in %	PPW in %	NPW in %
Sardanelli, 2007	278	18	59	-	77 (50 - 92)	-
Kuhl, 2005	618	12	42	-	29	-
Kuhl, 2005	529	43	33	97	28	-
Leach, 2005	649	35	40 (24 - 58)	93 (92 - 95)	10 (5,8 - 17)	99 (98-99)
Warner, 2004	236	22	38	100	83	97
Kriege, 2004	1909	45	33	95	48	-

In Klammern sind die 95 % Konfidenzintervalle angegeben, wenn sie sich aus den zitierten Arbeiten ermitteln ließen. (Sardanelli et al. 2007, Kuhl et al. 2005b, Kuhl et al. 2005a, Leach et al. 2005, Warner et al. 2004, Kriege et al. 2004), modifizierte Tabelle nach (Kuhl 2007a).

In der vorliegenden Arbeit waren sowohl Sensitivität, negativer prädiktiver Wert als auch diagnostische Genauigkeit der Röntgenmammographie signifikant kleiner als die der MR-Mammographie. Ebenfalls deutlich unterschieden sich die jeweils ermittelten *likelihood-ratios* im Sinne der klaren Überlegenheit der MR-Mammographie. Spezifität und positiver prädiktiver Wert unterschieden sich dagegen nicht signifikant in den beiden Bildgebungsmodalitäten. Als ein signifikanter Unterschied wurde das Fehlen einer Überschneidung der 95 %-Konfidenzintervalle der jeweiligen Werte betrachtet. Es wurden demnach MR-mammographisch weniger falsch negative Befunde und eine ähnliche Proportion an falsch positiven Befunden abgegeben, für die genauen Zahlen siehe Tabelle 17 auf Seite 36. Die Studien der Tabellen 35 und 36

geben zwar niedrigere Spezifitäten für die MR-Mammographie als für die Röntgenmammographie an, jedoch sind nur in einem Fall die Konfidenzintervalle beweisend für einen tatsächlichen Unterschied.

Die angegebenen Sensitivitäten der konventionellen Röntgenmammographie liegen sämtlich unterhalb der in der vorliegenden Arbeit ermittelten (64,8%). In zwei Fällen überschneiden sich jedoch die 95 %-Konfidenzintervalle, beziehungsweise umfasst das Konfidenzintervall den beschriebenen Wert (Sardanelli et al. 2007, Leach et al. 2005). Im Vergleich zu den in 7.2.1 aufgeführten Werten sind sie ebenfalls vergleichsweise niedrig. Als Grund hierfür kommt primär die Wahl des Studienkollektivs in Betracht. Da die Populationen in den zitierten Arbeiten unter einem erhöhten kumulativen Lebenszeitrisiko für Malignome der Mamma standen, erfolgte ihre Untersuchung im Mittel weit früher als die durchschnittliche Mammographie-Screening-Untersuchung (Warner 2008). Die Dichte des Brustdrüsenparenchyms ist abhängig vom Lebensalter, im Sinne einer stattfindenden Altersinvolution (Russo und Russo 2004). Oben wurde bereits dargelegt, dass die diagnostische Genauigkeit der Röntgenmammographie abhängig von eben der Parenchymdichte und damit auch direkt von dem Lebensalter des Patienten ist. Vermutlich kommt die Verteilung der Parenchymdichte auf die einzelnen Patienten in der vorliegenden Arbeit der Screeningsituation näher als in den zitierten Arbeiten. Tabelle 14 auf Seite 33 zeigt, dass 44 % der untersuchten Mammae über ein weniger dichtes Drüsenparenchym verfügten (ACR-Klasse 1 und 2), wie es gehäuft bei älteren Patienten auftritt.

### 7.3 Intervallkarzinome und falsch negative Befunde in der MRM

Um die Studienpopulation weiter auf ihre Vergleichbarkeit mit der Screeningpopulation zu untersuchen, ist die Angabe von Inzidenzraten für maligne intramammäre Neoplasien sinnvoll. Inzidenz bezeichne die Anzahl der Neuerkrankungen pro 1000 Individuen pro Jahr. In der vorliegenden Arbeit wurden initial 129 maligne Neoplasien diagnostiziert, die sich auf 116 Patienten verteilten. Nicht zu diesen 116 Patienten zählen die drei primär neu diagnostizierten Malignome sowie die vier Rezidive des weiteren klinischen Verlaufs, siehe 6.6.1. Für primäre Intervallkarzinome gilt folgendes: sie verteilten sich auf 226 (342-116) initial gesunde Patienten, die eine Follow-Up-Zeit von 119 Wochen (=2,3 Jahre; Median) hatten. Es wurden pro Jahr also 1,3 maligne Neoplasien pro 226 Patienten und damit 5,8 pro 1000 Patienten diagnostiziert. Die Angaben der Inzidenz in publizierten großen Screeningpopulationen schwanken

zwischen 0,9 und 2,2 pro 1000 Patienten pro Jahr (Frisell et al. 1987, Moss und Blanks 1998, Woodman et al. 1995). Es erkrankten also in einem nach Tabelle 8 (siehe Seite 24) vorselektierten Patientenkollektiv pro Jahr mehr als doppelt so viele Patienten neu an einem Mammakarzinom als in der Screeningpopulation. Zumindest ist dieser Unterschied nicht altersbedingt, die zitierten Werte aus den Screeningstudien beziehen sich auf die Kohorte der 54- bis 63-jährigen. Wie in 5.1.1 dargelegt, betrug das mittlere Alter zum Zeitpunkt der initialen MRT 54,4 Jahre, im Durchschnitt wurde mit 56,7 Jahren erneut untersucht.

In 6.6.2 wurde die Zusammensetzung der vier falsch negativen Befunde der initialen MR-Untersuchung vorgestellt. Bei drei dieser Malignome handelte es sich um einen ILC- sowie zwei DCIS-Herde, die jeweils weniger als drei Millimeter im Durchmesser maßen. Der Argumentation von 3.3 und 3.4.2 folgend, erhebt die MR-Mammographie keinen Anspruch derartige Läsionen mit Sicherheit detektieren zu können. Die Diagnostik von *in-situ*-Karzinomen mittels der MRM ist in ihrer Genauigkeit noch nicht abschließend bewertet. Zahlreiche ältere Studien geben hier deutlich niedrigere Werte für die Sensitivität der MR-Mammographie als für die korrespondierende Röntgenmammographie an (Orel et al. 1997, Boetes et al. 1997). Aktuelle Publikationen zeigen das Gegenteil (Neubauer et al. 2003, Vag et al. 2008).

DCIS-Läsionen, die ausschließlich in der konventionellen Röntgenmammographie sichtbar sind, wurden als signifikant kleiner beschrieben im Vergleich zu solchen, die ausschließlich oder auch mit der MR-Untersuchung detektierbar sind (Kuhl et al. 2007). Der Anteil von DCIS-Läsionen unter den falsch negativen Befunden in der MR-Mammographie wird als überproportional hoch angegeben (Wurdinger et al. 2001). Insgesamt folgen die falsch negativen Befunde der vorliegenden Arbeit diesen zusammengekommenen Charakteristika.

## 7.4 Genauigkeit der Bildgebung hinsichtlich der Tumorausbreitung

### 7.4.1 Tumordurchmesser

Das Ziel einer diagnostischen Prozedur ist es, dass sich, im Falle eines positiven Befundes, das nachfolgende Behandlungsschema ändert und optimiert wird. Je größer eine detektierte intramammäre Läsion ist, desto höher ist die Wahrscheinlichkeit, dass es sich um ein Malignom handelt (Lieberman et al. 2006). Mit zunehmender Größe eines Malignoms steigt ebenfalls die Wahrscheinlichkeit der Fernmetastasierung

(Thames et al. 1999), wobei diese das entscheidende Kriterium zur Verkürzung der Überlebensrate von unter Brustkrebs leidenden Patienten ist (Guarneri und Conte 2004). Die Vorhersage der tatsächlichen Tumorgröße anhand des bildgebenden Verfahrens ist demnach von enormer klinischer Bedeutung.

Dass sowohl die Messungen der Röntgenmammographie als auch der MR-Mammographie eine ausreichend gute Korrelation ( $r > 0,7$ ) mit den realen Tumorausdehnungen haben (Echevarria et al. 2006), kann bestätigt werden. Siehe hierzu Tabelle 19 auf Seite 38. Insgesamt existierten bezogen auf die 95 %-Konfidenzintervalle keine signifikanten Unterschiede zwischen den einzelnen Korrelationskoeffizienten. Die Vorhersage des tatsächlichen Tumordurchmessers mittels der MRM von unbehandelten Malignomen gilt in der Literatur als exzellent, es werden Korrelationskoeffizienten zwischen 0,96 (Röntgenmammographie:  $r = 0,76$ ) und 0,98 angegeben (Röntgenmammographie:  $r = 0,45$ ) (Davis et al. 1996, Mumtaz et al. 1997). Bei chemotherapeutisch behandelten Malignomen verliert die MR-Mammographie jedoch an Aussagekraft ( $r$  zwischen 0,43 und 0,65) bezüglich der Vorhersage des Durchmessers (Wasser et al. 2003, Kim et al. 2007). In der vorliegenden Arbeit gingen sowohl behandelte als auch unbehandelte Ergebnisse der Bildgebung in die Korrelation ein. Auch wurden die Befunde der Bildgebung nicht gemeinsam von Radiologen und Pathologen hinsichtlich ihrer Genauigkeit geprüft, wie es in prospektiven Studien zur Tumorausdehnung möglich ist (Echevarria et al. 2006). Als einschränkend muss weiterhin betrachtet werden, dass der maximale Tumordurchmesser im Falle von invasiven Malignomen im pathologischen Referenzbefund stets nur die invasive Komponente beschreibt, nicht jedoch eventuell assoziierte *in-situ*-Komponenten (Greene 2002). Die MR-Mammographie beurteilt die allen Läsionen eigene, zusätzlich entstandene Perfusion. Auf diese Weise mögen oftmals deutlich höhere Werte entstehen, als die histopathologische Referenz angibt. Ein signifikanter Unterschied zwischen den Größen von Malignomen die mittels konventioneller Röntgenmammographie detektiert werden konnten und solchen, zu denen alleine die MRM eine Aussage machte, bestand nicht. Siehe hierzu auch 6.7.

#### 7.4.2 Multifokalität und Multizentrität

Zahlreiche Studien zeigen die Überlegenheit der MR-Mammographie gegenüber der Röntgenmammographie, eine multifokale oder multizentrische Karzinomausbreitung zu detektieren (Orel et al. 1995, Mumtaz et al. 1997, Drew et al. 1999, Fischer et al. 1999). Es werden Korrelationskoeffizienten von bis zu 0,99 für die Aussage der Mul-

tifokalität beziehungsweise Multizentrität der MR-Bildgebung mit dem pathologischen Befund angegeben (Echevarria et al. 2006).

Jedoch wurde in den wenigsten Fällen eine histopathologische Untersuchung der kompletten Mamma als Referenzstandard für die MR-Mammographie verwendet (Sardanelli et al. 2004). Dies führt zwangsläufig dazu, dass der Anteil falsch positiver Foci-Befunde steigt. In einem Kollektiv bioptisch gesicherter Karzinome wurden Werte von 53 % für monofokale, 29 % für multifokale und 18 % für multizentrische Erkrankungen angegeben. (Sardanelli et al. 2004). In der vorliegenden Arbeit zeigte der pathologische Befund hierfür 63 %, 27 % sowie 10 %, siehe Tabelle 21 auf Seite 39. Der positive prädiktive Wert für monofokale Erkrankungen betrug MR-mammographisch 91,3 % (42 von 46) und röntgenmammographisch 81,1 % (30 von 37). Die Sensitivität betrug hingegen lediglich 60,8 % (42 von 69) respektive röntgenmammographisch 88,2 % (30 von 34, siehe Tabelle 20).

Demnach wurden zu 39,2 % der monofokalen Befunde in der MRT zusätzliche, falsch positive Läsionen beschrieben. Insgesamt wurden MR-mammographisch 4,9 % der Befunde unterbewertet (disseminiertere Tumorausdehnung in der histopathologischen Aufarbeitung als in der Bildgebung, 4 + 1 von 109, röntgenmammographisch 14,6 %) und 37,6 % überbewertet (röntgenmammographisch 8,3 %). Verwendet man dagegen die pathologische Aufarbeitung der kompletten Mamma als Referenz, werden MR-mammographisch 30,3 % (röntgenmammographisch 19,2 %) der Befunde überbewertet sowie 19,2 % (röntgenmammographisch 30,3 %) unterbewertet (Sardanelli et al. 2004). Aus dieser Studie und der vorliegenden Arbeit wird klar, dass die MRT eindeutig mehr maligne intramammäre Läsionen im Falle eines disseminierten Wachstums erkennt. Diese Genauigkeit geht jedoch einher mit einer höheren Anzahl falsch positiver Läsionsbefunde. Für die vorliegende Arbeit wurde keine dedizierte histopathologische Aufarbeitung durchgeführt, sondern die Ergebnisse der klinischen Routinediagnostik verwendet. Die größere Rate an zusätzlichen, falsch positiven Herdbefunden mag zeigen, dass ein Anteil prinzipiell vorhandener Informationen in der klinischen Routine verloren geht, mit denen einige der falsch positiven Befunde hätten vermieden werden können. Kritisch muss in der vorliegenden Arbeit die geringe Anzahl der Fälle betrachtet werden, in denen sowohl die Aussagen der Bildgebung als auch der Pathologie vorhanden waren (röntgenmammographisch  $n = 48$ , siehe Tabelle 20 auf Seite 39). Dies kann zu einer verzerrten Darstellung der wahren Verhältnisse geführt haben.

## 7.5 Die Reevaluation der MR-Mammographien

### 7.5.1 Veränderte diagnostische Genauigkeit

Aus Tabelle 26 auf Seite 43 geht hervor, dass die Reevaluation der MR-Mammographien die signifikant genaueste Methode war, maligne Herdbefunde von benignen zu unterscheiden (diagnostische Genauigkeit der Reevaluation: 91,4 %, im Vergleich zu der originären MR-Mammographie: 81,3 % sowie der Röntgenmammographie: 70,6 %). Im Unterschied zu den in 7.2.3 dargelegten Verhältnissen unterscheiden sich die letzten beiden Befundungsmodalitäten diesbezüglich nicht mehr signifikant. So ist die originäre MR-Mammographie zwar sensitiver, jedoch weniger spezifisch als die Röntgenmammographie, betrachtet man lediglich die histologisch verifizierten Herde. Als Gründe können die folgenden angenommen werden:

Erstens ist die Zahl der Untersuchungen auf die histologisch verifizierten Herde reduziert worden. Damit ist die Anzahl der richtig negativen Untersuchungen im Sinne eindeutig benigner Befunde die keiner histopathologischen Abklärung bedurften stark verringert. Vergleich man die Tabellen 11 und 15 mit den Tabellen 25 und 24 ist auffällig, dass sich die absolute Zahl falscher Befunde in der MR-Mammographie nicht änderte. Dies war jedoch der Fall bei der konventionellen Röntgenmammographie. Die Schlussfolgerung daraus ist, dass die Röntgenmammographie einen Anteil falsch positiver Befunde generiert, der mittels der MR-Untersuchung vermieden werden kann. Gleichzeitig gibt es im Gegenzug eine Anzahl falsch positiver Befunde, die allein durch die MRM verursacht wurde.

Zweitens wurden in der Befundung der originären MR-Mammographien objektivierbare und damit vermeidbare Fehleinschätzungen getroffen. Dies folgt aus der sehr guten Übereinstimmung der beiden reevaluierenden Befunder und ihrer überlegenen diagnostischen Genauigkeit.

Bezüglich der falsch negativen MR-Befunde lässt sich folgendes anmerken: Die deutlich höhere Spezifität der Reevaluation ging einher mit einer geringgradig, nicht signifikanten Verminderung der Sensitivität. Übersehen wurde zusätzlich ein DCIS des Malignitätsgrades 1, ein Stromasarkom sowie ein Malignomrest nach Chemotherapie, siehe 6.8.2. Da die Reevaluation ohne Kenntnis klinischer Angaben erfolgte und die Genauigkeit der MR-Untersuchung postchemotherapeutisch eingeschränkt ist (Wasser et al. 2003, Rieber et al. 1997), muss zumindest dieser letzte falsch negative Befund mit Vorsicht betrachtet werden.

### 7.5.2 Objektivierete Gründe für falsch positive MR-Befunde

In 6.8.4 wurde beschrieben, dass ein überproportional hoher Anteil der falsch positiven MRT-Befunde non-mass-Läsionen betraf. Die Charakterisierung der non-mass-Läsionen folgte in der Reevaluation streng den BIRADS-Kriterien für MR-Mammographie (Tardivon et al. 2007). Die beiden reevaluierenden Befunder waren mit den BIRADS-Kriterien und deren (mindestens tendenziellen) Aussagen vertraut.

Insgesamt unterschied sich die Verteilung der Kontrastmittelanreicherungen bei non-mass- und mass-Läsionen auf maligne und benigne Herde. Das Washout-Phänomen gilt als starker Prädiktor für die Malignität von intramammären Läsionen (Kaiser 1992, Boetes et al. 1994, Kuhl et al. 1999), wie es auch in Abbildung 11 im Falle der mass-Läsionen verdeutlicht wird. In den reevaluierten non-mass-Läsionen trat es mit einer Ausnahme jedoch überhaupt nicht auf. Abbildung 11 führt weiterhin zu der Aussage, dass lediglich eine kontinuierliche Anreicherung des Kontrastmittels einen Hinweis auf die Gutartigkeit einer non-mass-Läsion gibt. Aus den Tabellen 33 und 34 geht als einziger klarer Prädiktor für eine benigne Läsion ein Anreicherungsmuster vom *stippled*-Typ hervor (siehe Abbildung 5 auf Seite 18). Dies geht konform mit den beiden bisher explizit zu non-mass-Anreicherungen publizierten Arbeiten (Sakamoto et al. 2008, Tozaki und Fukuda 2006). Sechs von 14 falsch positiven Befunden unter den non-mass-Läsionen zeigten eine Anreicherung vom Typ *stippled*, siehe Tabelle 34 auf Seite 47. Unter Kenntnis der Literatur sowie der Ergebnisse der vorliegenden Arbeit sind solche Befunde daher mit großer Sicherheit benigne, eine maligne Bewertung sollte mit Vorsicht betrachtet werden.

Die BIRADS-Kriterien raten, Befunde mit einem Anreicherungsmuster vom *clumped*-Typ zumindest als suspekt zu bezeichnen (Tardivon et al. 2007). Tendenzuell, wenn auch nicht statistisch signifikant, stützen dies die beiden zitierten Arbeiten. In der vorliegenden Arbeit dagegen verteilte sich das Anreicherungsmuster gleichmäßig auf richtig positive (maligne) und falsch positive (und damit benigne) Befunde.

Für mass-Läsionen lässt sich aus den Tabellen 30 bis 32 die Aussage gewinnen, dass ein Großteil der falsch positiven Befunde (58,8%), rein morphologisch betrachtet, glatt berandet war. Eine glatte Berandung in Abwesenheit von einem Washout-Phänomen ist ein vielfach beschriebener, enorm starker Prädiktor für die Gutartigkeit eines Herdbefundes (Nunes et al. 1997, Kinkel et al. 2000, Tozaki et al. 2006). Damit lässt sich feststellen, dass ein Teil der falsch positiven Befunde hätte vermieden werden können, wenn sich der Untersucher strikter an die BIRADS-Kriterien gehalten hätte. Der berechtigte Einwand gegen dieses Vorgehen, die Senkung der Anzahl

falsch positiver Befunde gehe zwangsläufig einher mit neuen falsch negativen kann relativiert werden. In der klinischen Diagnostik achten geschulte Befunder von MR-Mammographien neben den BIRADS-Kriterien auf eine Vielzahl anderer Zeichen, die entweder für Bösartigkeit oder Gutartigkeit sprechen. Im Verbund mit diesen steigt die diagnostische Genauigkeit, eine glatte Berandung wird nicht mit absoluter Sicherheit als benigne eingestuft (Kaiser 2007). So waren in der vorliegenden Arbeit 3,8 % ( $n = 4$ ) der Malignome glatt berandet. Zwei Läsionen entsprachen einem intramammären Non-Hodgkin-Lymphom, welche aufgrund anderer Deskriptoren als der der typischen Mammakarzinome eindeutig als maligne zu erkennen waren. Die anderen beiden glatt berandeten Befunde waren ein sechs Millimeter messendes nicht näher klassifiziertes Malignom sowie ein invasiv lobuläres Karzinom. Schließt man die Lymphome aus der Aufstellung in Tabelle 31 auf Seite 46 aus, so verbleiben lediglich 1,9 % der Malignome in der Gruppe der glatt berandeten Läsionen. Die Aussagekraft einer glatten Berandung im Sinne der Gutartigkeit nimmt für primäre Mammakarzinome damit weiter zu.

6.8.4 lässt den folgenden Schluss zu: Je geringer der Durchmesser einer mass-Läsion ist, desto höher ist die Wahrscheinlichkeit für einen falsch positiven Befund. Der mittlere Durchmesser tatsächlich maligner Läsionen (23,6 mm) unterschied sich signifikant von dem mittleren Durchmesser falsch positiver Läsionen (16,9 mm).

## 8 Schlussfolgerungen

Die MR-Mammographie ist der Röntgenmammographie diagnostisch eindeutig überlegen. Die vorliegende Arbeit ist die erste, die ihr Patientenkollektiv konsekutiv aus der klinischen Anwendung rekrutiert und über ein akzeptables Loss-To-Follow-Up verfügt. Die Aussagen über die überlegene Sensitivität der MR-Mammographie, wie sie an anders selektierten Studienkohorten ermittelt wurden, können damit weiter gestützt werden.

Dichtes Brustdrüsenparenchym vermindert die diagnostische Genauigkeit der Röntgenmammographie enorm. In solchen Fällen sollte aufgrund der niedrigen Sensitivität von dieser Untersuchung als diagnostische Methode abgesehen werden und die MR-Mammographie Verwendung finden. Da sich dichtes Brustdrüsenparenchym bevorzugt bei jüngeren Patienten findet, ist der Gedanke über ein nach Alter gestaffeltes Diagnoseverfahren sinnvoll.

Bei der Befundung der MR-Mammographie müssen röntgenmammographisch mit BIRADS 3 bezeichnete Fälle mit besonderer Aufmerksamkeit betrachtet werden. Per definitionem sollte die Wahrscheinlichkeit für ein Malignom unter 2 % liegen. Erhalten diese Patienten jedoch eine Überweisung für eine MRM, steigt die Wahrscheinlichkeit eines Malignoms auf 23,9 % (siehe 6.3). Ebenfalls deutlich erhöht ist in einem Patientenkollektiv, welches gemäß der Studienkohorte der vorliegenden Arbeit rekrutiert wurde, die Wahrscheinlichkeit der Krebsneuerkrankung. Dies ist bei der Diskussion zur Festlegung von Zeitabständen zwischen einzelnen Screeninguntersuchungen von großem Wert.

Je kleiner der Durchmesser einer mittels der MR-Mammographie detektierten masseläsion ist, desto wahrscheinlicher wird ein falsch positiver Befund. Die Bestimmung verlässlicher Prädiktoren für die Gutartigkeit solcher Läsionen ist deshalb vonnöten. Die BIRADS-Kriterien bieten ein solches Klassifizierungslexikon. Die konsequente Anwendung dieser Kriterien kann, wie gezeigt wurde, einen Anteil der falsch positiven Befunde verhindern. Trotzdem ist eine ständige Prüfung und Weiterentwicklung der Kriterien anzustreben, um ihre Genauigkeit weiter zu verbessern. Die ermittelten positiven prädiktiven Werte für die Angabe eines eventuellen multifokalen oder multizentrischen Tumorwachstums und ihr Unterschied zu denen einer gezielt darauf ausgerichteten Studie lassen auf einen Informationsverlust im Rahmen der klinischen Routinediagnostik schließen.

Die vorliegende Arbeit zeigt erstmalig, dass ein überproportional großer Anteil an

falsch positiven Befunden non-mass-Läsionen betrifft. Die genauere Klassifizierung und die sich jeweils ableitende diagnostische Konsequenz sollten in Zukunft Gegenstand weiterer Studien sein, um die Unsicherheit im Umgang mit non-mass-Läsionen zu verringern. Auch die genauen histologischen Unterschiede zwischen mass- und non-mass-Läsionen bedürfen einer weiteren Abklärung.

Insgesamt kann damit die in 4 formulierte Nullhypothese verworfen werden. Es besteht ein signifikanter Unterschied in der diagnostischen Genauigkeit von Röntgenmammographie und MR-Mammographie zugunsten letzterer. Abschließend soll das Paradoxon, dass die MR-Mammographie aktuell keine anerkannte Screeningmethode ist (nicht einmal im Fall von dichtem Brustdrüsenparenchym), in Form einer logisch inkonsistenten Trias gezeigt werden.

- Erklärtes Ziel des Arztes ist es, dem einzelnen Menschen bestmöglich zu helfen (bestmögliche Hilfe definiert sich durch die Anwendung wissenschaftlich-empirisch gewonnener Erkenntnisse).
- Eingesetzt werden sollte zum Screening für Brustkrebs generell die konventionelle Röntgenmammographie.
- Die MR-Mammographie ist der Röntgenmammographie hinsichtlich der diagnostischen Genauigkeit überlegen.

Es ist nicht möglich, alle drei Punkte der Trias zu akzeptieren ohne in einen Widerspruch zu geraten. Da aktuell einzig die Röntgenmammographie als Mittel der Wahl zum Screening vorgeschlagen wird (siehe 3.4), müssen die Befürworter dieses Vorgehens Punkt eins oder drei abstreiten. Dies gestaltet sich für Punkt eins aus ethischen Gründen schwierig, für die Verifikation von Punkt drei existiert mittlerweile eine nahezu zwingende Studienlage. Das pauschale Argument der unzureichenden Verfügbarkeit der MR-Mammographie sollte aufgrund dieser Erwägungen kritisch geprüft werden.

## Literaturverzeichnis

- Albert U, Altland H, Duda V, Engel J, Geraedts M, Heywang-Köbrunner S, Hölzel D, Kalbheim E, Koller M, König K, et al. 2008. Kurzfassung der aktualisierten Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland 2008. *Geburtsh Frauenheilk*, 68:251–61.
- Albert U, Schreer I. 2008. Prävention und Früherkennung von Brustkrebs. *Der Onkologe*, 14(5):454–60.
- Altman D, Bland J. 1994. Diagnostic tests 2: predictive values. *British Medical Journal*(International ed), 309(6947).
- Bedrosian I, Mick R, Orel S, Schnall M, Reynolds C, Spitz F, Callans L, Buzby G, Rosato E, Fraker D, et al. 2003. Changes in the surgical management of patients with breast carcinoma based on preoperative magnetic resonance imaging. *Cancer*, 98(3):468–73.
- Berg W, Gutierrez L, NessAiver M, Carter W, Bhargavan M, Lewis R, Ioffe O. 2004. Diagnostic Accuracy of Mammography, Clinical Examination, US, and MR Imaging in Preoperative Assessment of Breast Cancer 1. *Radiology*, 233(3):830–49.
- Bertz J, Hentschel S, Hundsdörfer G, Kaatsch P, Katalinic A, Lehnert M, Schön D, Stegmaier C, Ziegler H. 2008. Krebs in Deutschland, Häufigkeiten und Trends. Arbeitsgemeinschaft Bevölkerungsbezogener Krebsregister in Deutschland in Zusammenarbeit mit dem Robert Koch Institut Saarbrücken.
- Bisgard K, Folsom A, Hong C, Sellers T. 1994. Mortality and Cancer Rates in Nonrespondents to a Prospective Study of Older Women: 5-Year Follow-up. *American Journal of Epidemiology*, 139(10):990–1000.
- Bland M. 2000. An introduction to medical statistics. Oxford University Press Oxford, UK.
- Boetes C, Barentsz J, Mus R, van der Sluis R, van Erning L, Hendriks J, Holland R, Ruys S. 1994. MR characterization of suspicious breast lesions with a gadolinium-enhanced TurboFLASH subtraction technique. *Radiology*, 193(3):777–81.

- Boetes C, Strijk S, Holland R, Barentsz J, Van Der Sluis R, Ruijs J. 1997. False-negative MR imaging of malignant breast tumors. *European Radiology*, 7(8):1231–34.
- Britton A, Murray D, Bulstrode C, McPherson K, Denham R. 1995. Loss to follow-up: does it matter? *Lancet*, 345(8963):1511–2.
- Brnic Z, Hebrang A. 2001. Breast compression and radiation dose in two different mammographic oblique projections: 45 and 60 degrees. *Eur J Radiol*, 40(1):10–5.
- Burhenne H, Burhenne L, Goldberg F, Hislop T, Worth A, Rebbeck P, Kan L. 1994. Interval breast cancers in the Screening Mammography Program of British Columbia: analysis and classification. *American Journal of Roentgenology*, 162(5):1067–71.
- Bydder G, Young I. 1985. MR imaging: clinical use of the inversion recovery sequence. *J Comput Assist Tomogr*, 9(4):659–75.
- Canguilhem G. 1989. *The normal and the pathological*. Zone Books New York.
- Chevalier M, Morán P, Ten J, Soto J, Cepeda T, Vañó E. 2004. Patient dose in digital mammography. *Medical Physics*, 31:2471–9.
- Cochrane A, Holland W. 1971. Validation of Screening Procedures. *British Medical Bulletin*, 27(1):3–8.
- Collins L, Schafer J, Kam C. 2001. A comparison of inclusive and restrictive strategies in modern missing data procedures: New approaches to missing data. *Psychological methods*, 6(4):330–51.
- Davis P, Staiger M, Harris K, Ganott M, Klementaviciene J, McCarty K, Tobon H. 1996. Breast cancer measurements with magnetic resonance imaging, ultrasonography, and mammography. *Breast Cancer Research and Treatment*, 37(1):1–9.
- van Diest P. 1999. Ductal carcinoma in situ in breast carcinogenesis. *Journal of pathology*, 187(4):383–84.
- Drew P, Chatterjee S, Turnbull L, Read J, Carleton P, Fox J, Monson J, Kerin M. 1999. Dynamic Contrast Enhanced Magnetic Resonance Imaging of the Breast Is Superior to Triple Assessment for the Pre-Operative Detection of Multifocal Breast Cancer. *Annals of Surgical Oncology*, 6(6):599–603.

- Duijm L, Guit G, Zaat J, Koomen A, Willebrand D. 1997. Sensitivity, specificity and predictive values of breast imaging in the detection of cancer. *British Journal of Cancer*, 76(3):377–81.
- Echevarria J, Martín M, Saiz A, Imaz I, Fernández-Ruanova B, Martín D, López-Ruiz J. 2006. Overall Breast Density in MR Mammography: Diagnostic and Therapeutic Implications in Breast Cancer. *Journal of Computer Assisted Tomography*, 30(1):140–7.
- Elmore J, Barton M, Moceri V, Polk S, Arena P, Fletcher S. 1998. Ten-Year Risk of False Positive Screening Mammograms and Clinical Breast Examinations. *New England Journal of Medicine*, 338(16):1089–96.
- Fischer U. 2008. Röntgenmammographie. *Der Radiologe*, 48(8):785–98.
- Fischer U, Kopka L, Grabbe E. 1999. Breast Carcinoma: Effect of Preoperative Contrast-enhanced MR Imaging on the Therapeutic Approach 1. *Radiology*, 213(3):881–8.
- Fletcher S, Elmore J. 2003. Mammographic Screening for Breast Cancer. *New England Journal of Medicine*, 348(17):1672–80.
- Folkman J. 1995. Clinical Applications of Research on Angiogenesis. *New England Journal of Medicine*, 333(26):1757–63.
- von Fournier D, Abel U, Spratt J, Spratt J, Anton H. 1994. Wachstumsgeschwindigkeit des Mammakarzinoms, Bedeutung für Früherkennung und Therapieeffekte. *Geburtshilfe und Frauenheilkunde*, 54(5):286–90.
- Frisell J, Eklund G, Hellström L, Somell A. 1987. Analysis of interval breast carcinomas in a randomized screening trial in Stockholm. *Breast Cancer Research and Treatment*, 9(3):219–25.
- Gardner M, Altman D. 1986. Confidence intervals rather than P values: estimation rather than hypothesis testing. *British Medical Journal (Clinical research ed)*, 292(6522):746–50.
- Gates T. 2001. Screening for cancer: evaluating the evidence. *Am Fam Physician*, 63(3):513–22.

- Gilbert F, Astley S, Gillan M, Agbaje O, Wallis M, James J, Boggis C, Duffy S. 2008. Single Reading with Computer-Aided Detection for Screening Mammography. *The New England Journal of Medicine*, 359(16):1675–84.
- Greene F. 2002. *AJCC cancer staging manual*. Springer New York.
- Guarneri V, Conte P. 2004. The curability of breast cancer and the treatment of advanced disease. *European Journal of Nuclear Medicine and Molecular Imaging*, 31:149–61.
- Heywang S, Hahn D, Schmidt H, Krischke I, Eiermann W, Bassermann R, Lissner J. 1986. MR imaging of the breast using gadolinium-DTPA. *Journal of computer assisted tomography*, 10(2):199–204.
- Heywang-Köbrunner S, Schlegel A, Beck R, Wendt T, Kellner W, Lommatzsch B, Untch M, Nathrath W. 1993. Contrast-Enhanced MRI of the Breast after Limited Surgery and Radiation Therapy. *Journal of Computer Assisted Tomography*, 17(6):891–900.
- Heywang-Köbrunner S, Viehweg P, Heinig A, Küchler C. 1997. Contrast-enhanced MRI of the breast: accuracy, value, controversies, solutions. *European journal of radiology*, 24(2):94–108.
- Hojilla C, Wood G, Khokha R. 2008. Metalloproteinases as common effectors of inflammation and extracellular matrix breakdown in breast cancer. *Breast Cancer Research*, 10(2):205.
- Holmgren L, O'Reilly M, Folkman J, et al. 1995. Dormancy of micrometastases: Balanced proliferation and apoptosis in the presence of angiogenesis suppression. *Nature Medicine*, 1(2):149–53.
- Houserkova D, Prasad S, Svach I, Kucerova L, Duskova M, Bucil J, Sisola I, Zlamalova N, Svebisova H. 2008. The value of dynamic contrast enhanced breast MRI in mammographically detected bi-rads 5 microcalcifications. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*, 152(1):107–15.
- Ikeda D, Hylton N, Kinkel K, Hochman M, Kuhl C, Kaiser W, Weinreb J, Smazal S, Degani H, Viehweg P, et al. 2001. Development, Standardization, and Testing of a Lexicon for Reporting Contrast-Enhanced Breast Magnetic Resonance Imaging Studies. *Journal of Magnetic Resonance Imaging*, 13:889–95.

- Jacobs M, Barker P, Bluemke D, Maranto C, Arnold C, Herskovits E, Bhujwala Z. 2003. Benign and Malignant Breast Lesions: Diagnosis with Multiparametric MR Imaging 1. *Radiology*, 229(1):225–32.
- Kaiser W. 1992. MRM promises earlier breast cancer diagnosis. *Diagn Imaging (San Franc)*, 14(9):88–93.
- Kaiser W. 2007. *Signs in MR-Mammography*. Springer.
- Kaiser W, Zeitler E. 1989. MR imaging of the breast: fast imaging sequences with and without Gd-DTPA. Preliminary observations. *Radiology*, 170(3 Pt 1):681–6.
- Kawashima H, Matsui O, Suzuki M, Kadoya M, Tawara M, Nonomura A, Noguchi M, Takashima T. 2000. Breast Cancer in Dense Breast: Detection With Contrast-Enhanced Dynamic MR Imaging. *Imaging*, 11:233–43.
- Kerlikowske K, Grady D, Barclay J, Sickles E, Ernster V. 1996. Effect of age, breast density, and family history on the sensitivity of first screening mammography. *JAMA*, 276(1):33–8.
- Kerlikowske K, Grady D, Rubin S, Sandrock C, Ernster V. 1995. Efficacy of screening mammography. A meta-analysis. *JAMA*, 273(2):149–54.
- Kim H, Im Y, Han B, Choi N, Lee J, Kim J, Choi Y, Ahn J, Nam S, Park Y, et al. 2007. Accuracy of MRI for estimating residual tumor size after neoadjuvant chemotherapy in locally advanced breast cancer: relation to response patterns on MRI. *Acta Oncol*, 46(7):996–1003.
- Kinkel K, Helbich T, Esserman L, Barclay J, Schwerin E, Sickles E, Hylton N. 2000. Dynamic High-Spatial-Resolution MR Imaging of Suspicious Breast Lesions Diagnostic Criteria and Interobserver Variability. *American Journal of Roentgenology*, 175(1):35–43.
- Kolb T, Lichy J, Newhouse J. 2002. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27825 patient evaluations. *Radiology*, 225(1):165–75.
- Kriege M, Brekelmans C, Boetes C, Besnard P, Zonderland H, Obdeijn I, Manoliu R, Kok T, Peterse H, Tilanus-Linthorst M, et al. 2004. Efficacy of MRI and Mammography for Breast-Cancer Screening in Women with a Familial or Genetic Predisposition. *New England Journal of Medicine*, 351(5):427–37.

- Kristman V, Manno M, Côté P. 2004. Loss to Follow-Up in Cohort Studies: How Much is Too Much? *European Journal of Epidemiology*, 19(8):751–60.
- Kuhl C. 2007a. Current Status of Breast MR Imaging\* Part 2. Clinical Applications. *Radiology*, 244(3):672–91.
- Kuhl C. 2007b. The Current Status of Breast MR Imaging\* Part I. Choice of Technique, Image Interpretation, Diagnostic Accuracy, and Transfer to Clinical Practice. *Radiology*, 244(2):356–78.
- Kuhl C, Mielcareck P, Klaschik S, Leutner C, Wardelmann E, Gieseke J, Schild H. 1999. Dynamic Breast MR Imaging: Are Signal Intensity Time Course Data Useful for Differential Diagnosis of Enhancing Lesions? *Radiology*, 211(1):101–10.
- Kuhl C, Schrading S, Bieling H, Wardelmann E, Leutner C, Koenig R, Kuhn W, Schild H. 2007. MRI for diagnosis of pure ductal carcinoma in situ: a prospective observational study. *The Lancet*, 370(9586):485–92.
- Kuhl C, Schrading S, Leutner C, Morakkabati-Spitz N, Wardelmann E, Fimmers R, Kuhn W, Schild H. 2005a. Mammography, Breast Ultrasound, and Magnetic Resonance Imaging for Surveillance of Women at High Familial Risk for Breast Cancer. *Journal of Clinical Oncology*, 23(33):8469–76.
- Kuhl C, Schrading S, Weigel S, Nüssele-Kügele K, Sittek H, Arand B, Morakkabati N, Leutner C, Tombach B, Nordhoff D, et al. 2005b. Die EVA-Studie: Evaluierung der Leistungsfähigkeit diagnostischer Verfahren(Mammographie, Sonographie, MRT) zur sekundären und tertiären Prävention des familiären Mammakarzinoms- Zwischenergebnisse nach der ersten Hälfte der Förderungsperiode. 177(6):818–27.
- Kurtz J, Jacquemier J, Amalric R, Brandone H, Ayme Y, Hans D, Bressac C, Spitalier J. 1990. Breast-conserving therapy for macroscopically multiple cancers. *Annals of Surgery*, 212(1):38–44.
- Landis J, Koch G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–74.
- Leach M, et al. 2005. Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: A prospective multicentre cohort study (MARIBS). *Journal Watch(General)*:1769–78.

- Liberman L, Mason G, Morris E, Dershaw D. 2006. Does Size Matter? Positive Predictive Value of MRI-Detected Breast Lesions as a Function of Lesion Size. *American Journal of Roentgenology*, 186(2):426–30.
- Mandelson M, Oestreicher N, Porter P, White D, Finder C, Taplin S, White E. 2000. Breast Density as a Predictor of Mammographic Detection: Comparison of Interval- and Screen-Detected Cancers. *Journal of the National Cancer Institute*, 92(13):1081–87.
- Masroor I. 2008. Effectiveness of assigning BI-RADS category-3 to Breast lesion with respect to follow-up. *J Coll Physicians Surg Pak*, 18(4):209–12.
- Michaelson J, Silverstein M, Wyatt J, Weber G, Moore R, Halpern E, Kopans D, Hughes K. 2002. Predicting the survival of patients with breast carcinoma using tumor size. *Cancer*, 95(4):713–23.
- Monninkhof E, Elias S, Vlems F, van der Tweel I, Schuit A, Voskuil D, van Leeuwen F. 2007. Physical Activity and Breast Cancer: A Systematic Review. *Epidemiology*, 18(1):137–57.
- Morgan M, Cooke M, McCarthy G. 2005. Microcalcifications Associated with Breast Cancer: An Epiphenomenon or Biologically Significant Feature of Selected Tumors? *Journal of Mammary Gland Biology and Neoplasia*, 10(2):181–87.
- Moss S, Blanks R. 1998. Calculating appropriate target cancer detection rates and expected interval cancer rates for the UK NHS Breast Screening Programme. Interval Cancer Working Group. *Journal of Epidemiology & Community Health*, 52(2):111–15.
- Mumtaz H, Hall-Craggs M, Davidson T, Walmsley K, Thurell W, Kissin M, Taylor I. 1997. Staging of symptomatic primary breast cancer with MR imaging. *American Journal of Roentgenology*, 169(2):417–24.
- Neubauer H, Li M, Kuehne-Heid R, Schneider A, Kaiser W. 2003. High grade and non-high grade ductal carcinoma in situ on dynamic MR mammography: characteristic findings for signal increase and morphological pattern of enhancement. *British Journal of Radiology*, 76(901):3–12.

- Noel A, Munaut C, Nusgens B, Lapiere C, Foidart J. 1993. Different mechanisms of extracellular matrix remodeling by fibroblasts in response to human mammary neoplastic cells. *Invasion Metastasis*, 13(2):72–81.
- Nothacker M, Lelgemann M, Giersiepen K, Weinbrenner S. 2007. zur S-3-Leitlinie Brustkrebsfrüherkennung in Deutschland. Band 31. ed Version 1.00. Berlin: Ärztliches Zentrum für Qualität in der Medizin (ÄZQ).
- Nunes L, Schnall M, Orel S, Hochman M, Langlotz C, Reynolds C, Torosian M. 1997. Breast MR imaging: interpretation model. *Radiology*, 202(3):833–41.
- Orel S, Kay N, Reynolds C, Sullivan D. 1999. BI-RADS Categorization As a Predictor of Malignancy 1. *Radiology*, 211(3):845–50.
- Orel S, Mendonca M, Reynolds C, Schnall M, Solin L, Sullivan D. 1997. MR imaging of ductal carcinoma in situ. *Radiology*, 202(2):413–20.
- Orel S, Schnall M, Powell C, Hochman M, Solin L, Fowble B, Torosian M, Rosato E. 1995. Staging of suspected breast cancer: effect of MR imaging and MR-guided biopsy. *Radiology*, 196(1):115–22.
- Pisano E, Hendrick R, Yaffe M, Baum J, Acharyya S, Cormack J, Hanna L, Conant E, Fajardo L, Bassett L, et al. 2008. Diagnostic Accuracy of Digital versus Film Mammography: Exploratory Analysis of Selected Population Subgroups in DMIST. *Radiology*, 246(2):376–83.
- Rieber A, Brambs H, Gabelmann A, Heilmann V, Kreienberg R, Kühn T. 2002. Breast MRI for monitoring response of primary breast cancer to neo-adjuvant chemotherapy. *European Radiology*, 12(7):1711–19.
- Rieber A, Zeitler H, Rosenthal H, Gorich J, Kreienberg R, Brambs H, Tomczak R. 1997. MRI of breast cancer: influence of chemotherapy on sensitivity. *British Journal of Radiology*, 70(833):452–8.
- Robertson C. 1993. A Private Breast Imaging Practice: Medical Audit of 25,788 Screening. *Radiology*, 187:75–9.
- Russo J, Russo I. 2004. Development of the human breast. *Maturitas*, 49(1):2–15.

- Sakamoto N, Tozaki M, Higa K, Tsunoda Y, Ogawa T, Abe S, Ozaki S, Sakamoto M, Tsuruhara T, Kawano N, et al. 2008. Categorization of non-mass-like breast lesions detected by MRI. *Breast Cancer*, 15(3):1–6.
- Saleh A, Kurz K, Mödder U. 2005. BI-RADS update: Mammographie, Brustultraschall und Kernspinnmammographie BI-RADS update: Mammography, Breast Ultrasound and Magnetic Resonance Imaging of the Breast. *Radiologie up2date*, 5:285–312.
- Santen R, Mansel R. 2005. Benign Breast Disorders. *New England Journal of Medicine*, 353(3):275–85.
- Sardanelli F, Giuseppetti G, Panizza P, Bazzocchi M, Fausto A, Simonetti G, Lattanzio V, Del Maschio A. 2004. Sensitivity of MRI Versus Mammography for Detecting Foci of Multifocal, Multicentric Breast Cancer in Fatty and Dense Breasts Using the Whole-Breast Pathologic Examination as a Gold Standard. *American Journal of Roentgenology*, 183(4):1149–57.
- Sardanelli F, Podo F, D’Agnolo G, Verdecchia A, Santaquilani M, Musumeci R, Trecate G, Manoukian S, Morassut S, de Giacomo C, et al. 2007. Multicenter Comparative Multimodality Surveillance of Women at Genetic-Familial High Risk for Breast Cancer (HIBCRIT Study): Interim Results. *Radiology*, 242(3):698–715.
- Schorn C, Fischer U, Luftner-Nagel S, Westerhof J, Grabbe E. 1999. MRI of the breast in patients with metastatic disease of unknown primary. *European Radiology*, 9(3):470–73.
- Sickles E. 1991. Periodic Mammographic Follow-up of Probably Benign Lesions: Results in 3,184 Consecutive Cases. *Radiology*, 179:463–68.
- Sickles E. 1999. Probably Benign Breast Lesions: When Should Follow-up Be Recommended and What Is the Optimal Follow-up Protocol? 1. *Radiology*, 213(1):11–4.
- Sims A. 1973. Importance of a high tracing-rate in long-term medical follow-up studies. *Lancet*, 2(7826):433–5.
- Smigal C, Jemal A, Ward E, Cokkinides V, Smith R, Howe H, Thun M. 2006. Trends in Breast Cancer by Race and Ethnicity: Update 2006. *CA: A Cancer Journal for Clinicians*, 56(3):168–83.

- Smith-Bindman R, Chu P, Miglioretti D, Quale C, Rosenberg R, Cutter G, Geller B, Bacchetti P, Sickles E, Kerlikowske K. 2005. Physician Predictors of Mammographic Accuracy. *JNCI*, 97(5):358–67.
- Tardivon A, Athanasiou A, Thibault F, El Khoury C. 2007. Breast imaging and reporting data system (BIRADS): Magnetic resonance imaging. *European Journal of Radiology*, 61(2):212–15.
- Thames H, Buchholz T, Smith C. 1999. Frequency of First Metastatic Events in Breast Cancer: Implications for Sequencing of Systemic and Local-Regional Treatment. *Journal of Clinical Oncology*, 17(9):2649–58.
- Tozaki M, Fukuda K. 2006. High-Spatial-Resolution MRI of Non-Masslike Breast Lesions: Interpretation Model Based on BI-RADS MRI Descriptors. *American Journal of Roentgenology*, 187(2):330–37.
- Tozaki M, Igarashi T, Fukuda K. 2006. Positive and Negative Predictive Values of BI-RADS®-MRI Descriptors for Focal Breast Masses. *Magnetic Resonance in Medical Sciences*, 5(1):7–15.
- Twisk J, de Vente W. 2002. Attrition in longitudinal studies How to deal with missing data. *Journal of Clinical Epidemiology*, 55(4):329–37.
- Vag T, Baltzer P, Renz D, Pfeleiderer S, Gajda M, Camara O, Kaiser W. 2008. Diagnosis of ductal carcinoma in situ using contrast-enhanced magnetic resonance mammography compared with conventional mammography. *Clinical imaging*, 32(6):438–42.
- Van Goethem M, Tjalma W, Schelfout K, Verslegers I, Biltjes I, Parizel P. 2006. Magnetic resonance imaging in breast cancer. *European Journal of Surgical Oncology*, 32(9):901–910.
- Varas X, Leborgne J, Leborgne F, Mezzera J, Jaumandreu S, Leborgne F. 2002. Revisiting the Mammographic Follow-Up of BI-RADS Category 3 Lesions. *American Journal of Roentgenology*, 179(3):691–95.
- de Visser K, Coussens L. 2006. The Inflammatory Tumor Microenvironment and Its Impact on Cancer Development. *Contributions to Microbiology*, 13(R):118–37.

- Warner E. 2008. The Role of Magnetic Resonance Imaging in Screening Women at High Risk of Breast Cancer. *Topics in Magnetic Resonance Imaging*, 19(3):163–9.
- Warner E, Plewes D, Hill K, Causer P, Zubovits J, Jong R, Cutrara M, DeBoer G, Yaffe M, Messner S, et al. 2004. Surveillance of BRCA1 and BRCA2 Mutation Carriers With Magnetic Resonance Imaging, Ultrasound, Mammography, and Clinical Breast Examination. *JAMA*, 292(11):1317–25.
- Wasser K, Sinn H, Fink C, Klein S, Junkermann H, Lüdemann H, Zuna I, Delorme S. 2003. Accuracy of tumor size measurement in breast cancer using MRI is influenced by histological regression induced by neoadjuvant chemotherapy. *European Radiology*, 13(6):1213–23.
- Weidner N. 1995. Intratumor microvessel density as a prognostic factor in cancer. *The American Journal of Pathology*, 147(1):9–19.
- Woodman C, Threlfall A, Boggis C, Prior P. 1995. Is the three year breast screening interval too long? Occurrence of interval cancers in NHS breast screening programme's north western region. *British Medical Journal*, 310(6974):224–26.
- Wurdinger S, Kamprath S, Eschrich D, Schneider A, Kaiser W. 2001. False-negative findings of malignant breast lesions on preoperative magnetic resonance mammography. *The Breast*, 10(2):131–39.
- Youden W. 1950. Index for rating diagnostic tests. *Cancer*, 3(1):32–5.

## 9 Anhang

### 9.1 Danksagung

Mein Dank gilt Herrn Professor Dr. W. A. Kaiser für die Überlassung des Themas, die zahlreichen Diskussionen und das immerwährende Interesse am Fortgang der Arbeit.

Besonders danken möchte ich Herrn Dr. P. A. T. Baltzer für die engagierte und stimulierende Betreuung der Arbeit und für die lehrreiche Zeit insgesamt am Institut für Interventionelle und Diagnostische Radiologie.

Nicht zuletzt gilt mein Dank den zahlreichen niedergelassenen Kollegen, ohne die das niedrige Loss-To-Follow-Up nicht möglich gewesen wäre.

## 9.2 Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass mir die Promotionsordnung der Medizinischen Fakultät der Friedrich-Schiller-Universität bekannt ist,

ich die Dissertation selbst angefertigt habe und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben sind,

mich folgende Personen bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts unterstützt haben: Prof. Dr. med. Dipl.-Chem. W. A. Kaiser und Dr. med. P. A. T. Baltzer,

die Hilfe eines Promotionsberaters nicht in Anspruch genommen wurde und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen,

dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe und

dass ich die gleiche, eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung nicht bei einer anderen Hochschule als Dissertation eingereicht habe.

Jena, 18.04.2012

Matthias Benndorf