

FABÍOLA PROCACI KESTELMAN

**VARIAÇÃO INTEROBSERVADOR E VALOR PREDITIVO POSITIVO
DA CATEGORIA 4 DO BI-RADS® PARA MAMOGRAFIA**

Dissertação de Mestrado

**ORIENTADOR: Prof. Dr. GUSTAVO ANTONIO DE SOUZA
CO-ORIENTADOR: Prof. Dr. LUÍS OTÁVIO ZANATTA SARIAN**

**Unicamp
2008**

FABÍOLA PROCACI KESTELMAN

**VARIAÇÃO INTEROBSERVADOR E VALOR PREDITIVO POSITIVO
DA CATEGORIA 4 DO BI-RADS® PARA MAMOGRAFIA**

Dissertação de Mestrado apresentada à
Pós-Graduação da Faculdade de Ciências
Médicas da Universidade Estadual de
Campinas para obtenção do Título de
Mestre em Tocoginecologia, área de
Tocoginecologia

ORIENTADOR: Prof. Dr. GUSTAVO ANTONIO DE SOUZA
CO-ORIENTADOR: Prof. Dr. LUÍS OTÁVIO ZANATTA SARIAN

Unicamp
2008

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA FACULDADE DE CIÊNCIAS MÉDICAS
UNICAMP**

Bibliotecário: Sandra Lúcia Pereira – CRB-8ª / 6044

K489v Kestelman, Fabíola Procaci
Variação inter-observador e valor preditivo positivo da
categoria 4 do BL-RADS® para mamografia / Fabíola
Procaci Kestelman. Campinas, SP: [s.n.], 2008.

Orientadores: Gustavo Antonio de Souza, Luís Otávio
Sarian
Dissertação (Mestrado) Universidade Estadual de
Campinas. Faculdade de Ciências Médicas.

1. Mama - cancer. 2. Mamografia. I. Souza, Gustavo
Antonio de. II. Sarian, Luís Otávio. III. Universidade
Estadual de Campinas. Faculdade de Ciências Médicas.
IV. Título.

Título em inglês: **Reader variability and predictive positive value for mammography
BI-RADS® category 4**

Keywords: • **Breast neoplasm**
 • **Mammography**

Titulação: **Mestre em Tocoginecologia**
Área de concentração: **Tocoginecologia**

Banca examinadora: **Prof. Dr. Gustavo Antonio de Souza**
 Prof. Dr. Cláudio Kemp
 Profa. Dra. Júlia Yoriko Shinzato

Data da defesa: **15 – 05 – 2008**

BANCA EXAMINADORA DA DISSERTAÇÃO DE MESTRADO

Aluna: FABÍOLA PROCACI KESTELMAN

Orientador: Prof. Dr. GUSTAVO ANTONIO DE SOUZA

Co-Orientador: Prof. Dr. LUÍS OTÁVIO ZANATTA SARIAN

Membros:

1.

2.

3.

Curso de Pós-Graduação em Tocoginecologia da Faculdade
de Ciências Médicas da Universidade Estadual de Campinas

Data: 15/05/2008

200814949

Dedico este trabalho...

*Ao meu amor, marido, amigo, Eduardo Kestelman,
por ter sido cúmplice desta conquista e
exemplo por toda minha história na medicina.*

*Aos meus filhos, Bernardo e Felipe,
pela compreensão na minha ausência,
por serem a razão de todas as minhas conquistas.*

*Aos meus pais, Procaci e Maria Inês, sobretudo pelo amor e
também pela orientação e exemplo de retidão e perseverança.
Pelo alicerce proporcionado na minha vida.*

*Aos meus irmãos, Rafael e Cintia,
que completam a família acolhedora que eu sempre tive.*

*Não há flor sem raízes,
nem frutos sem chão fértil.*

Agradecimentos

Ao Prof. Dr. Gustavo Antonio de Souza, pelo exemplo de competência na medicina, pesquisa científica e na orientação deste estudo.

Ao Prof. Dr. Luís Otávio Zanatta Sarian um grande mestre, pela dedicação e disponibilidade necessárias para a realização deste trabalho.

À Profa. Dra. Sophie Françoise Mauricette Derchain, pela amizade e orientações valiosas neste processo de aprendizagem.

As Dras. Vivianne Aquilhera Rolin de Freitas e Gabriela Martins, pelas horas de trabalho, dedicadas com competência e carinho a este projeto.

À Prof. Ellyete de Oliveira Canella, por ter despertado o prazer pela radiologia e por ter sido incondicionalmente minha amiga.

À amiga Dra. Marcela Balara pela amizade e diponibilidade que tanto necessitei.

À amiga Lilian Scheinkman, pelo apoio, incentivo e amizade sempre presentes.

Ao Prof. Dr. Luiz Cláudio Thuler pelo carinho com que auxiliou este projeto junto ao INCA e à Unicamp.

A amiga Renata Rizzo Matheus pela ajuda na organização do projeto e pela amizade neste processo.

Aos médicos da Clínica Felipe Mattoso pela amizade, companheirismo e apoio, incentivo, amizade e disposição para ajudar sempre.

A equipe de trabalho da clínica Felipe Mattoso e do INCA, sempre presente para socorrer nos momentos delicados.

Agradecimentos Institucionais

Ao Instituto Nacional de Câncer (INCA) pelo financiamento parcial deste estudo.

Ao Departamento de Tocoginecologia da Faculdade de Ciências Médicas da Unicamp, em especial aos professores participantes do Mestrado, pela viabilização do curso de pós-graduação do qual este trabalho é fruto.

Sumário

Símbolos, Siglas e Abreviaturas.....	ix
Resumo	x
Summary	xii
1. Introdução.....	14
2. Objetivos.....	22
2.1. Objetivo geral.....	22
2.2. Objetivos específicos.....	22
3. Publicação	23
4. Conclusões.....	47
5. Referências Bibliográficas	48
6. Anexos.....	52
6.1. Anexo 1 – Ficha para coleta de dados – Radiologista 1	52
6.2. Anexo 2 –Ficha para coleta de dados – Radiologista 2	53
6.3. Anexo 3 –Ficha para coleta de dados – Radiologista 3	54
6.4. Anexo 4 - Carta de aprovação do projeto CP – SPC – DPQ	55
6.5. Anexo 5 - Cartas de aprovação do projeto no CEP – INCA.....	56

Símbolos, Siglas e Abreviaturas

BI-RADS® – *Breast Imaging Reporting and Data System*

CAISM – Centro de Atenção Integral à Saúde da Mulher

CEP – Comitê de Ética em Pesquisa

CONEP – Comissão Nacional de Ética em Pesquisa

DTG – Departamento de Tocoginecologia

FCM – Faculdade de Ciências Médicas

INCA – Instituto Nacional de Câncer

MS – Ministério da Saúde

OMS – Organização Mundial da Saúde

Unicamp – Universidade Estadual de Campinas

VPP – Valor Preditivo Positivo

WHO – *World Health Organization*

Resumo

Objetivo: Determinar o valor preditivo positivo das categorias 4A, 4B e 4C da 4ª edição do BI-RADS e determinar a variação interobservador desta classificação em lesões submetidas à biópsia cirúrgica. **Métodos:** Este estudo retrospectivo foi conduzido no Instituto Nacional de Câncer (INCA), Rio de Janeiro, Brasil. Os casos para análise foram selecionados no banco de dados do setor de radiologia, entre mulheres submetidas à marcação pré-cirúrgica para biópsia cirúrgica entre janeiro de 1999 e dezembro de 2002, que foram previamente classificadas como BI-RADS 4. As mamografias foram revisadas por três radiologistas especializados em radiologia mamária. Os avaliadores classificaram as lesões considerando a probabilidade de malignidade, dentro da categoria 4, em 4A, 4B e 4C. Para a análise estatística, foi realizado um consenso a partir dos resultados dos três radiologistas. A variação interobservador foi calculada utilizando o índice de *kappa* ponderado. **Resultados:** De 775 lesões, o resultado consensual mostrou 235 (30,3%) mamografias classificadas como 4A, 338 (43,6%) como 4B e o restante 161 (20,8%) como 4C. Os radiologistas não obtiveram consenso em 41 (5,3%) lesões. A concordância geral entre os avaliadores foi moderada ($\kappa=0,44$), variando entre $\kappa=0,38$ considerando os radiologistas dois e três e $\kappa=0,54$ considerando os

radiologistas um e dois. A concordância variou de mínima (categoria 4B, $k= 0,09$) a pobre, considerando as categorias 4A e 4C. A proporção de lesões nos grupos 4A, 4B e 4C foi similar e o VPP foi progressivo (23,0%, 47,6% e 72,7%, respectivamente). **Conclusão:** O VPP das categorias 4A, 4B and 4C foi progressivo e a concordância entre os observadores foi moderada.

Summary

Purpose: To determine the predictive positive value of BI-RADS 4th edition categories 4A, 4B and 4C and determine the inter-observer variability for lesion on which biopsies had been performed. **Methods:** This cross-sectional study was conducted at the National Institute of Cancer (INCA), Rio de Janeiro, Brazil. Cases for analysis were selected from the medical records among women who were submitted to breast needle localization and diagnostic surgical biopsy, from January 1999 to December 2002, and which had been previously categorized as BIRADS 4. The mammograms were reviewed by three senior breast imaging radiologists. Readers provided a BI-RADS final assessment category to indicate the probability of malignancy in category 4. For statistical calculations, a “consensus” variable was created using the outputs from the three assessors. Interobserver variability for each subcategory was determined using the weighted *kappa* statistic. **Results:** Of 775 lesions, “consensual” assessment rendered 235 (30.3%) mammograms classified as 4A, 338 (43.6%) as 4B and the remainder 161 (20.8%) as 4C. Radiologists did not reach a consensus in 41 (5.3%) exams. Overall agreement between raters was moderate ($\kappa=0.44$), ranging from $\kappa=0.38$ between raters two and three to $\kappa=0.54$ between radiologists one and two.

Agreement between raters ranged from slight (category 4B, $k=0.09$) to fair, considering categories 4A and 4C. In BI-RADS 4B and 4C groups, the proportion of each abnormality in the whole group was similar to that in 4A, but the PPV were higher. **Conclusions:** The PPV for subcategories 4A, 4B and 4C was progressive and the agreement between readers was moderate.

Key words: Breast cancer, mammography, BI-RADS

1. Introdução

O câncer de mama é a principal causa de morte por neoplasia entre mulheres no Brasil e corresponde ao segundo mais incidente, de acordo com as estatísticas do Instituto Nacional de Câncer (Brasil, 2006). A redução da taxa de mortalidade depende da detecção precoce, do planejamento terapêutico adequado e do uso da mamografia anual como método de rastreamento (Newman, 2003).

A mamografia possui alta sensibilidade para detectar câncer de mama clinicamente oculto. Uma revisão dos ensaios clínicos que avaliaram o desempenho do método mostrou que a sensibilidade variou entre 71% e 98% para o exame anual (Thuler, 2003). Porém, muitas lesões consideradas suspeitas, para as quais é indicado estudo histopatológico, correspondem a alterações benignas. Nos Estados Unidos, o valor preditivo positivo (VPP) das biópsias realizadas por achados mamográficos, ou seja, o total de lesões malignas diagnosticadas sobre o total de biópsias realizadas, varia entre 15% e 40% (Ciatto et al., 1987; Cyrllak, 1988; Hall et al., 1988). O custo e a morbidade das intervenções para o diagnóstico destas lesões são levados em conta para a confirmação do uso da mamografia como método de rastreamento (Heywang-Kobrunner, 2001).

Uma das dificuldades da análise do exame é que grande parte das lesões diagnosticadas não apresenta características patognomônicas. Knutzen e Gisvold (1993) estudaram a probabilidade de malignidade em várias categorias de lesões não-palpáveis detectadas na mamografia e observaram que se os critérios morfológicos destas lesões fossem considerados, a taxa de lesões malignas entre mulheres submetidas à biópsia poderia chegar a 40%.

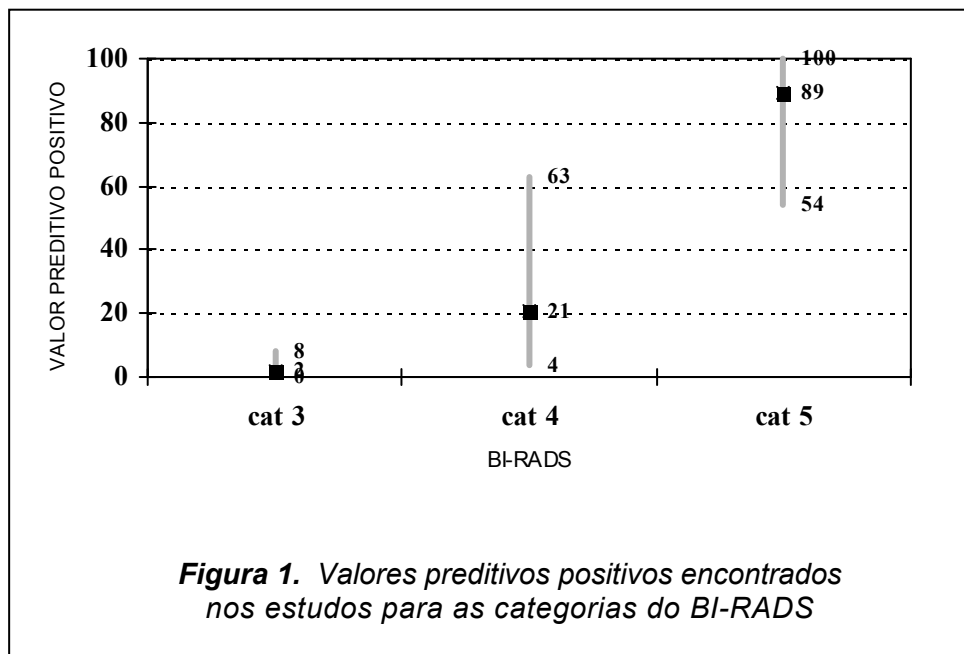
Considerando então a necessidade de padronização do laudo mamográfico e visando reduzir as discordâncias na interpretação da mamografia, o American College of Radiology publicou, em 1993, o Breast Imaging Reporting and Data System (BI-RADS) (American College of Radiology, 1993). Novas edições foram publicadas em 1995, 1998 e 2003 (American College of Radiology, 1995; American College of Radiology, 1998; American College of Radiology, 2004). De acordo com a quarta edição do BI-RADS® (American College of Radiology, 2004), os exames são classificados com base no grau de suspeição das lesões em: categoria 1 - sem achados positivos, categoria 2 - achados benignos, categoria 3 - achados provavelmente benignos, categoria 4 - achados suspeitos e categoria 5 - achados altamente suspeitos. Lesões que necessitam de avaliação adicional, por exemplo, com ultra-sonografia, são classificadas na categoria 0 e aquelas que têm diagnóstico histopatológico maligno previamente confirmado, na categoria 6.

Alguns estudos avaliam a capacidade para prever malignidade das categorias 3, 4 e 5 onde seriam classificadas as lesões com algum grau de suspeição. Uma maneira de verificar o desempenho de cada categoria do BI-

RADS® é analisando o resultado das lesões que foram submetidas à biópsia e calcular o VPP, tendo como padrão-ouro o resultado histopatológico.

Onze estudos determinaram os VPP das categorias 3, 4 e 5 (Bérubé et al., 1998; Liberman et al., 1998; Lacquement et al., 1999; Orel et al., 1999; Margolin et al., 2001; Ball et al., 2002; Travade et al., 2002; Mendez et al., 2004; Tan et al., 2004; Zonderland et al., 2004). As características de todos os estudos estão resumidas na Tabela 1.

Nos 11 estudos o VPP variou entre 0% e 8% para a categoria 3, sendo a mediana de 2%, entre 4% e 63% para a categoria 4, mediana de 21% e entre 54% e 100% na categoria 5 com mediana igual a 89% (Figura 1).



Três artigos mencionaram, além do VPP das categorias BI-RADS, as características morfológicas das lesões com maior associação com malignidade (Tabela 2). Bérubé et al (1998) e Liberman et al. (1998) avaliaram as características morfológicas das lesões como proposto na segunda edição do BI-RADS (American College of Radiology, 1995), incluindo para o estudo dos nódulos a avaliação da margem e da forma e para as calcificações a morfologia e a distribuição. Mendez et al. (2004) dividiram as lesões em microcalcificações, densidade assimétrica, nódulo circunscrito, espiculado ou com microcalcificações e densidade assimétrica com microcalcificações. Nos três estudos, o nódulo espiculado foi a lesão com maior VPP. Liberman et al. (1998) verificaram que, com relação aos nódulos, os às microcalcificações foram morfologia linear e distribuição segmentar e critérios com maior VPP foram margem espiculada e forma irregular. Com relação linear. Bérubé et al. (1998) identificaram que os critérios morfológicos com maior VPP foram os nódulos densos, nódulos com margens espiculadas e microcalcificações lineares.

Um fator importante para a análise dos VPP das categorias é a limitação do uso do BI-RADS, que pode estar associado a deficiências da própria classificação ou do treinamento dos radiologistas que vão utilizá-la (Liberman e Menell, 2002). Berg et al. (2000) avaliaram a variabilidade inter e intra-observador do uso da terminologia do BI-RADS. Cinco radiologistas com experiência em mamografia avaliaram 103 mamografias de rotina e 83 diagnósticas. A taxa de concordância entre os radiologistas (método estatístico de *Kappa*) variou entre 0,16 e 0,77 para os diversos achados mamográficos,

mostrando uma moderada variabilidade e de 0,37 para as categorias do BI-RADS que significa uma concordância discreta. Orel et al. (1999), que analisaram achados de mamografias realizadas em instituições diferentes, observaram que algumas pacientes que eram encaminhadas para realização de biópsia apresentavam lesões benignas, classificadas na categoria 2 e, portanto, não teriam indicação de biópsia. Essa discrepância foi associada com a variação interobservador, tanto para descrever as lesões quanto para recomendar a biópsia, e também foi relacionada com variações na experiência dos radiologistas que avaliaram as mamografias. Bérubé et al. (1998) associaram o baixo VPP que obtiveram para a categoria 4 com o fato da nomenclatura do BI-RADS ser descritiva e pouco específica.

Ao contrário de Bérubé et al. (1998), Zonderland et al. (2004) encontraram VPP maiores para todas as categorias em comparação com resultados encontrados em estudos americanos. Este achado foi atribuído ao fato de que nos estudos americanos há uma tendência em se obter um maior número de diagnósticos falsos positivos para reduzir o número de resultados falsos negativos, o que interfere na seleção das pacientes para investigação histopatológica. O mesmo achado foi encontrado em um estudo comparando mamografias de rotina realizadas nos Estados Unidos e no Reino Unido (Smith et al., 2003).

Com relação aos valores preditivos positivos das categorias, o BI-RADS sugere valores abaixo de 2% para a categoria 3 e acima de 95% para a categoria 5, sendo que cinco estudos encontraram valores acima do sugerido para a categoria 3 (Lacquement et al., 1999; Margolin et al., 2001; Mendez et al., 2004;

Tan et al., 2004; Zonderland et al., 2004) e em nove artigos os valores obtidos foram mais baixos que o esperado para a categoria 5 (Bérubé et al., 1998; Liberman et al., 1998; Lacquement et al., 1999; Margolin et al., 2001; Tate et al., 2001; Ball et al., 2002; Travade et al., 2002; Mendez et al., 2004; Tan et al., 2004;).

Mesmo levando-se em conta estas variações, estes resultados confirmam uma significativa diferença no desempenho entre as categorias, sugerindo a utilidade da classificação. A dificuldade para a utilização do BI-RADSTM parece estar então relacionada com a divisão das lesões de acordo com as categorias, considerando os diferentes critérios morfológicos, o que tornaria a classificação menos subjetiva e com maior concordância inter e intra-observador.

Em 2003 o American College of Radiology publicou a 4^a edição do BI-RADS[®] propondo algumas alterações (American College of Radiology, 2003). Em resposta aos comentários feitos pelos usuários das edições anteriores e com o objetivo de dar informações sobre as mudanças da nova edição, uma nova seção foi inserida (Guidance Chapter).

A categoria 6 foi acrescentada nesta edição, destinada a lesões identificadas no exame que tenham sido previamente submetidas à biópsia e apresentam resultado histopatológico maligno.

De acordo com a definição de que a maioria das lesões que são consideradas suspeitas é classificada como categoria 4 e que esta categoria apresentaria um VPP bastante variável (entre 3% e 94%), a nova edição propõe também uma alteração na categoria 4. Foi sugerida uma subdivisão opcional

em 4A, 4B e 4C. A categoria 4A deve ser utilizada para as lesões que necessitem de investigação histopatológica, mas tenham baixo grau de suspeição. Neste caso não é esperado um laudo histopatológico maligno. Na categoria 4B estariam classificadas as lesões com grau de suspeição intermediário. Um estudo com resultado histopatológico benigno deve ser correlacionado com o achado radiológico. A categoria 4C inclui achados com alto grau de suspeição, porém que não são clássicos para malignidade. Um resultado maligno é esperado para esta categoria.

Esta divisão, entretanto, é opcional, sendo útil para auditorias e para orientar decisões de condutas. Esta divisão pode auxiliar o médico assistente e o patologista na correlação do resultado da biópsia com o achado mamográfico, por exemplo. Seria aceitável um resultado histopatológico benigno para lesões na categoria 4A, porém lesões da categoria 4C com resultado de biópsia benigno devem continuar a investigação diagnóstica. Sendo assim, a proposta desta classificação seria auxiliar na definição da conduta das lesões classificadas na categoria 4, antes e após a biópsia, sem o objetivo de contraindicar a investigação histopatológica das lesões incluídas nesta categoria.

Contudo, o BI-RADS® permite certa subjetividade para subdividir as lesões nas diversas categorias. As dúvidas sobre como classificar os exames criam um questionamento quanto à utilidade da subdivisão, pois pode não haver diferença significativa nos VPP de cada subcategoria conforme foi sugerido. Caso isso se confirme, a subdivisão seria desnecessária, levando ao aumento da subjetividade e do tempo para confecção do laudo mamográfico, sem benefício evidente.

Lazarus et al. (2006) avaliaram o VPP das categorias 4A, 4B e 4C e a variação interobservador e verificaram um VPP de 6%, 15% e 53% respectivamente, mostrando uma gradação para malignidade, de acordo as categorias. Entretanto, quando avaliaram a variação interobservador, encontraram uma concordância pobre, de 0,14, 0,16 e 0,28 para as categorias 4A, 4B e 4C respectivamente. A variação interobservador também foi avaliada por Ciatto et al. (2006), que verificaram 50 lesões - avaliadas por 12 observadores - que demonstraram uma concordância mínima, de 0,08, 0,07, 0,10, para as categorias 4A, 4B e 4C respectivamente.

Estes estudos mostram que as categorias 4A, 4B e 4C podem apresentar diferenças significativas na capacidade de prever malignidade e que pode haver uma concordância interobservador que justifica sua utilização. Partindo deste pressuposto, o que demonstraria a utilidade desta subdivisão, propusemos o atual estudo.

2. Objetivos

2.1. Objetivo geral

Determinar o valor preditivo positivo e verificar a variação interobservador das categorias 4A, 4B e 4C do BI-RADS® em mamografias de mulheres com lesões não palpáveis, tendo como padrão-ouro o estudo histopatológico.

2.2. Objetivos específicos

- Calcular o VPP das categorias 4A, 4B e 4C de lesões identificadas em mamografias de mulheres submetidas à biópsia cirúrgica, considerando a análise histopatológica como padrão-ouro.
- Verificar a variação interobservador das categorias 4A, 4B e 4C, considerando três radiologistas como observadores.

3. Publicação

> Date: Wed, 9 Apr 2008 11:26:03 -0400
> From: office@european-radiology.org
> To: fabkest@hotmail.com
> Subject: European Radiology - Manuscript ID ER-Apr-2008-004987
>
> 09-Apr-2008
>
>
> Dear Dr. Sarian
>
> Your manuscript entitled "Reader variability and predictive positive value of the sub-divisions of mammography BI-RADS category 4" has been successfully submitted online and is presently being given full consideration for publication in European Radiology.
>
> Your manuscript ID is ER-Apr-2008-004987.
>
> Please mention the above manuscript ID in all future correspondence or when calling the office for questions. If there are any changes in your postal address or e-mail address, please log in to Manuscript Central at <https://mc.manuscriptcentral.com/eurradiol> and edit your user information as appropriate.
>
> You can also view the status of your manuscript at any time by checking your Author Center after logging in to <https://mc.manuscriptcentral.com/eurradiol> .
>
> Thank you for submitting your manuscript to European Radiology.
>
> Yours sincerely
> With kind regards
>
> European Radiology Editorial Office

Reader variability and predictive positive value for mammography BI-RADS category 4

Fabiola Procaci Kestelman, MD (1)

Luis Otavio Sarian, MD, PhD (2)

Sophie Derchain, MD, PhD (2)

Vivianne Aguilera Rolim de Freitas (1)

Gabriela Martins (1)

Ellyete de Oliveira Canella (1)

Gustavo Antonio de Souza (2)

(1) Departamento of Radiology, Brazilian National Cancer Institute - INCA, Rio de Janeiro, Brazil.

(2) Department of Obstetrics and Gynecology, Universidade Estadual de Campinas (Unicamp), Campinas, Brazil.

Corresponding author: Luís Sarian

sarian@terra.com.br

Key words: breast cancer, mammography, BI-RADS

Reader variability and predictive positive value of the sub-divisions of mammography BI-RADS category 4

Objective: To determine the predictive positive value of BI-RADS 4th edition categories 4A, 4B and 4C and determine the inter-observer variability for lesion on which biopsies had been performed.

Methods: This cross-sectional study was conducted at the National Institute of Cancer (INCA), Rio de Janeiro, Brazil. Cases for analysis were selected from the medical records among women who were submitted to breast needle localization and diagnostic surgical biopsy, from January 1999 to December 2002, and which had been previously categorized as BIRADS 4. The mammograms were reviewed by three senior breast imaging radiologists. Readers provided a BI-RADS final assessment category to indicate the probability of malignancy in category 4. For statistical calculations, a “consensus” variable was created using the outputs from the three assessors. Interobserver variability for each subcategory was determined using the weighted *kappa* statistic.

Results: Of 775 lesions, “consensual” assessment rendered 235 (30.3%) mammograms classified as 4A, 338 (43.6%) as 4B and the remainder 161 (20.8%) as 4C. Radiologists did not reach a consensus in 41 (5.3%) exams. Overall agreement between raters was moderate ($\kappa=0.44$), ranging from $\kappa=0.38$ between raters two and three to $\kappa=0.54$ between radiologists one and two. Agreement between raters ranged from slight (category 4B, $k=0.09$) to fair, considering categories 4A and 4C. In BI-RADS 4B and 4C groups, the proportion of each abnormality in the whole group was similar to that in 4A, but the PPV were higher.

Conclusions: The PPV for subcategories 4A, 4B and 4C was progressive and the agreement between readers was moderate.

Introduction

There is general agreement that screening mammography reduces the rate of deaths among women with breast cancer who are 40 years old or older (1,2). In a meta-analysis of eight randomized trials, in women assigned to screening mammography, the mortality rate of was reduced in up to 35% among women aged 50-69 years, and of 15% to 20% among those aged 40-49. However, the major flaw of screening mammograms resides in the relatively high number of false-positive results, i.e. lesions considered amenable to surgical excision but with a benign final pathologic status (3, 4).

Besides the low specificity of the mammography, the variability in the terminology used in mammography reporting leads to inconsistent recommendations (5). To address concerns about terminology, the American College of Radiology created the Breast Imaging Reporting and Data System (BI-RADS®), which includes seven numbered assessment categories of 0-6 with associated management recommendations (6). In the 4th edition the lexicon suggested an optional subdivision of category 4 (4A, 4B, 4C) that is used for the majority of findings prompting breast interventional procedures. This subdivision allows a more meaningful practice audit and is an aid for clinicians and pathologists. The category 4A may be used for a finding needing intervention but with a low suspicion for malignancy. The category 4B includes those with intermediate suspicion and the category 4C includes findings of moderate concern, but not classic for malignancy (as in category 5) (6).

The variability in the application of the BI-RADS assessment categories for mammography has been studied to a large extent. In the studies with the most optimistic results, fair to moderate agreement was reported (7). As expected, the division of category 4 into three other subdivisions led to further variability and disagreement between radiologists. The current mammographic lexicon does not explicitly state which mammography features should be included in each subdivision and the studies evaluating these new categories are scarce. A recent study obtained fair agreement for category 4C ($\kappa = 0,26$), however there was poor agreement between observers for category 4A and 4B ($\kappa = 0,14$ and $0,16$ respectively) (8).

There are strict guidelines for the management of breast abnormalities detected with mammography. However, radiologists accustomed to examining mammograms notice that breast surgeons may feel uncomfortable with the current BI-RADS subdivision of the category 4, especially regarding the blurred differentiation between categories 3 and 4A. There is a current scarcity of large and well-conducted studies that examine the inter-observer variability of category 4 of the BI-RADS. Moreover, there is no current evidence confirming an upward trend in the positive predictive value of the sub-divisions of category 4 in actual practice. The purpose of the present study was to examine these important parameters in a large sample of over 770 mammography studies, thereby providing consistent understanding of the limitations posed by the current BI-RADS recommendations for mammography.

Subjects and methods

Patient Selection

This cross-sectional study was conducted at the National Institute of Cancer (INCA), Rio de Janeiro, Brazil. Institutional review board (IRB) approval was obtained. The IRB granted exemption of the informed consent form because of the retrospective and non-interventionist nature of the study.

Cases for analysis were selected from the medical records among women who were submitted to breast needle localization and diagnostic surgical biopsy, from January 1999 to December 2002. Of 1179 cases, 906 (76.8%) had been previously considered under category 4. A total of 131 cases were excluded, 49 because mammographic films were unavailable for review, 35 had only ultrasonographic presentation, 27 were considered category 2, 3, 5 or 6 and 20 were excluded because final pathologic diagnoses were not retrieved. The remaining 775 lesions in 678 patients comprised the study population. The median age for these patients was 55.7 years (90% central range 45-79 years).

Mammography

Bilateral mammography was performed with a dedicated mammography unit (Mammomat 3000 Siemens; or Alpha RT, Instrumentarium). Patients underwent initial craniocaudal and true lateral mammography, with magnification mammograms available in some cases. Radiographic assessment rendered 116 mass lesions,

407 calcifications, 98 asymmetries, 71 architectural distortions and 83 combined lesions (calcification with mass - 34, asymmetry with calcification – 32 and architectural distortion – 17).

Imaging Interpretation and Data Analysis

The mammograms were reviewed by three senior breast imaging radiologists. Readers provided a BI-RADS final assessment category to indicate the probability of malignancy in category 4. Observers were asked to review the films without comparison films. The observers were familiar with the sub-divisions of BI-RADS category 4, and all observers had at least three years of experience using the fourth edition of BI-RADS. They were trained using the lexicon “Guidance Chapter” orientation, but the decision of the subcategory was individual and subjective. For statistical calculations, a “consensus” variable was created using the outputs from the three assessors. The “consensus” diagnosis was determined when at least two readers agreed upon one of the sub-divisions of category 4 for one given mammogram. When the three readers rendered different diagnoses, the mammogram received the “complete disagreement” label. The nature of all lesions was confirmed with histopathologic results after surgical excision. Of the 775 lesions, 348 (44.9%) were malignant and 427 (55.1%) were benign. Malignant masses included infiltrating ductal carcinomas ($n = 194$), infiltrating lobular carcinomas ($n = 17$), ductal carcinoma in situ (DCIS) ($n = 117$) and infiltrative lobular and ductal carcinomas ($n = 3$). Benign lesions ($n = 427$) included fibrocystic changes, papillomas, radial scars, and fibroadenomas. For statistical purposes, all

invasive carcinomas and *in situ* ductal carcinoma were allotted to the malignant group and the remainders were classified as benign, following the recommendations issued by BI-RADS®.

Data was recorded in Open Office spreadsheets and statistical calculations were performed with the R statistical package (R Environment). Interobserver variability for each subcategory was determined using the weighted *kappa* statistic. The method for estimating an overall weighted *kappa* value in the case of multiple observers and multiple categories is based on the work of Landis and Koch (9). The guidelines of Landis and Koch were followed in interpreting κ values: 0.00–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.80–1.00, almost perfect agreement.

Results

“Consensual” assessment rendered 235 (30.3%) mammograms classified as 4A, 338 (43.6%) as 4B and the remainder 161 (20.8%) as 4C. Radiologists did not reach a consensus in 41 (5.3%) exams. Overall agreement between raters was moderate ($\kappa=0.44$), ranging from $\kappa=0.38$ between raters two and three to $\kappa=0.54$ between radiologists one and two (Table 1).

In the mammograms rendered as category 4A, calcifications, masses and asymmetries were the most common findings, comprising well over 85% of the exams. The PPV of these findings ranged from 14.8 (asymmetry) to 24.8

(calcification). Although rare in this group (4A), architectural distortions were associated with cancer in all cases, as a stand-alone abnormality or associated to calcifications. In BI-RADS 4B and 4C groups, the proportion of each abnormality in the whole group was quite similar to that in 4A, but the PPV were higher. The mammograms for which radiologists disagreed completely, the PPV of the abnormal findings stood between those for categories 4B and 4C (Table 2).

Agreement between raters ranged from slight (category 4B, $k = 0.09$) to fair, considering categories 4A and 4C (Table 3).

Discussion

The present study is an attempt to improve the overall knowledge on the PPV of the three sub-divisions of BI-RADS category 4, and the study is based on one of the largest samples examined in this regard so far. The PPV progressed steadily from category 4A to 4C, and our results may help determine the likelihood of a women having clinically significant breast disease when receiving a BI-RADS 4 diagnosis. The American College of Radiology BI-RADS was developed to provide a standardized reporting system for mammography, to minimize the variability in mammography interpretation, to maximize the accuracy of the screening by the influence in the number of false positive exams (6, 10), but confirmation of these goals may only be attained with large clinical studies.

The final assessment categories of the BI-RADS lexicon are useful to predict malignancy. A search for studies evaluating the PPV of categories 3, 4

and 5 disclosed 11 publications, all of them describing different conditions for mammographic interpretation. In those studies, the PPV ranged from 0% to 8%, 4% to 63% and 54% to 100% for categories 3, 4 and 5 respectively (11). Lazarus et al. (8) evaluated 94 mammographic lesions (32% malignant) and analyzed the PPV for lesions categorized as BI-RADS 4 and 5. In that study, the PPV were 6%, 15% and 53% for categories 4A, 4B and 4C respectively and 91% for category 5. Our study corroborates these previous findings, because the PPV improved from category 4A to 4C. However, there is an important difference between our findings and those of the previous studies: our PPV were generally higher when compared to that obtained from other samples. These differences may be attributed to the fact that INCA is an institution totally dedicated to the study and treatment of cancer, which is likely to favor a selection bias. The institution receives cases referred from basic health units nationwide, and these cases are much likely pre-selected.

Lee et al. (12) studied 150 lesions that were first examined with breast ultrasound, after which histological samples were obtained with ultrasound-guided breast core biopsy. In the ultrasound assessment, the PPV were 26%, 89% and 90% for categories 4A, 4B and 4C respectively, and 97% for category 5, confirming the upward trend for PPV in line with the BI-RADS ultrasound classification. Even though radiologists have been using the BI-RADS lexicon for decades, large variability may still be seen among radiology practitioners. Variability in mammographic interpretation had been reported even before the use of the BI-RADS lexicon, and this variability itself prompted ACR to design

the BI-RADS. Elmore et al. (5) published a study on 150 mammograms, from which 27 cases of cancer were diagnosed. Work up recommendation was highly variable ranging from 74% to 96% for women with cancer and 11% to 65% for women without cancer.

Since the introduction of the BI-RADS lexicon, observer variability has been re-evaluated by several authors (7, 13-15). These studies indicate that, even in the presence of a standardized lexicon, variability in mammographic reports still persists. Berg et al. (7) reported a highly variable lesion management, with a final assessment *kappa* of 0,37. Pijnappel et al. (15) reported moderate overall agreement for non-palpable lesions with microcalcification (*kappa* 0,54). Two studies evaluated the variability to categories 4A, 4B and 4C (8, 16). Ciatto et al. (16) published a study in which 50 lesions were tested by 12 radiologists and reported a slight agreement for category 4A, 4B and 4C (0.08, 0.07, 0.10 respectively). The overall agreement in that study was slight (0.20). Twelve dedicated breast radiologists were enrolled in this study without prior working knowledge of the classification system. They demonstrated that the variability was higher for the less experienced group (> 10 years: K range 0.39 –0.65; < 10 years range 0.26-0.66), although even expert radiologists might have limited consistency in using the BI-RADS assessment categories. The present study, however, has as a major limitation the fact that only one mammogram was used for rater assessment and an experienced radiologist before the start of the study had carried out the selection of this mammogram.

Some authors attempted to compare the performance of radiologists with large experience with breast images (8). In that study, five experienced breast radiologists, familiarized with the “Guidance Chapter” criteria for BI-RADS sub-categorization, evaluated 94 lesions. As Ciatto et al. (16), they showed fair agreement for the overall assessment category ($k = 0.28$). Nevertheless, the agreement for category 4C was fair ($k = 0.26$), and the agreement was only slight for categories 4A and 4B ($k = 0.14$ and 0.16 respectively). Comparing these figures to ours, the studies differed in that we had a lower agreement rate for category 4B. In our study, we found fair agreement for categories 4A and 4C (0.27 and 0.39), however the agreement for category 4B was only slight (0.09).

Our study showed higher overall agreement compared to these prior studies. Three reasons seem plausible to explain this fact: first, the present study was restricted to mammograms previously classified as category 4. Second, three radiologists rated the mammograms, whereas five radiologists were involved in the study by Lazarus et al. (8) and 12 in Ciatto et al. (16). And third and foremost, a higher agreement in our study can also be related to the prior extensive expertise of three radiologists involved in the study, very familiar with the “guidance chapter” in the fourth edition of the BI-RADS and with many years of practice in a reference national breast cancer center. Berg et al. (17) demonstrated that training has an immediate and remarkable impact on raters’ performance. They found that after a 1-day training session, expert consensus improved for features analysis and final assessment and this effect was maintained over 2-3 months.

The variability in interpretation of breast exams can also be demonstrated in ultrasonography. Lee et al. (12) demonstrated the inter-observer agreement for sonographic final assessment. They found good overall agreement, although the variability to the subcategories was high. The poorest agreement was found for category 4B (0.09), probably because only 4.8% of the lesions were in this category. In our study, we also found the poorest agreement for category 4B although the prevalence of lesions in this category was found to be higher (43.6%).

The agreement is higher in studies that combine US and mammography. Lee et al. (12) found moderate agreement to US descriptors, like shape and orientation ($k = 0.49$ and 0.56 respectively). Lazarus et al. (8) found high agreement to describe mass and calcification in mammography ($k = 0.84$ and 0.94 respectively). Moreover, they found moderate agreement to descriptors to ultrasound, such as lesion orientation ($k = 0.61$), shape ($k = 0.66$) and margin ($k = 0.40$). The variability concerning the lexicon also can be demonstrated by the analysis of the breast mammographic density. Ciatto et al. (18) showed that the categorization of breast density according to BI-RADS is feasible and consistency is reasonable between readers. The present study has a limitation concerning the study sample, comprised exclusively of lesions classified as BI-RADS 4, excluding those categorized as 3 or 5. This aspect of the study design may constitute an important selection bias, considering the fact that categories 4A and 4C are subject to down- and upstaging to categories 3 and 5, respectively.

The present study demonstrated progressive PPV from BI-RADS category 4A to 4C. Although the agreement for each category was slight to fair, the overall

agreement was moderate. Variability is inheriting in the practice of radiology and is not necessarily problematic, because extensive and continued training can possibly diminish it. It is important, however, to determine the necessity to stratify the lesions in these new categories with a focus on their impact on clinical decision-making. The present study provides reassurance to radiologists that they must seek for continuous training on BI-RADS, because even radiologists with extensive experience with mammography struggled to reach fair rates of agreement in category BI-RADS 4. Down- and upstaging of this category may have severe impacts on the management and outcome of women. Besides the clinical impact, insufficient training may negatively affect the overall treatment costs of mammographic abnormalities.

References

1. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002; 137(5 Part):347-60.
2. Institute of Medicine. Saving women's lives: integration and innovation: a framework for progress in early detection and diagnosis of breast cancer. Washington, D.C.: *National Academies Press*, 2005.
3. Hall FM, Storella JM, Silverstone DZ, Wyshak G. Nonpalpable breast lesions: recommendation for biopsy based on suspicion of carcinoma at mammography. *Radiology* 1988; 167(2):353-358.
4. Ciatto S, Cataliotti L, Distant V. Nonpalpable lesions detected with mammography: review of 512 consecutive cases. *Radiology* 1987; 165:99-102.
5. Elmore JG, Wells CK, Lee CH, Howard DH, Feinstein AR. Variability in Radiologists' interpretations of mammograms. *N Engl J Med* 1994; 331(22):1493-1499.
6. American College of Radiology. Breast imaging report and data system, breast imaging atlas. 4th ed. Reston, VA: American College of Radiology; 2003.
7. Berg WA, Campassi C, Langenberg P, Sexton MJ. Breast Imaging Reporting and Data System: inter- and intraobserver variability in feature analysis and final assessment. *AJR* 2000; 174:1769-1777.

8. Lazarus E, Mainiero MB, Schepps B, Koelliker SL, Livingston LS. BI-RADS Lexicon for US and Mammography: Interobserver Variability and Positive Predictive Value. *Radiology* 2006; 239(2):385-391.
9. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33(1):159 -174.
10. D'Orsi CJ. American College of Radiology mammography lexicon: an initial attempt to standardize terminology. *AJR* 1996; 166(4):779-780.
11. Kestelman FP, Souza GA, Thuler LC, Martins G, Freitas VAR, Canella EO. Valor Preditivo positivo das categories 3, 4 e 5. Revisão sistemática da literatura. **Radiol Bras** 2007; 40(3):173-77.
12. Lee HJ, Kim EK, Kim MJ et al. Observer variability of Breast Imaging Reporting and Data System (BI-RADS) for breast ultrasound. *Eur J Radiol*, 2008; 65(2):293-298.
13. Baker JA, Kornguth PJ, Floyd CE. Breast imaging reporting and data system standatdized mammography lexicon: observer variability in lesion description. *AJR* 1996; 166:773-778.
14. Kerlikowske K, Grady D, Barclay J, Frankel SD, Ominsky SD, Sickles EA. Variability and accuracy in mammographic interpretation using the American College of Radiology breast imaging reporting and data system. *J Natl Cancer Inst* 1998; 90(23): 1801-1809.

15. Pijnappel RM, Peeters PH, Hendriks JH, Mali WP. Reproducibility of mammographic classifications for non-palpable suspect lesions with microcalcifications. *Br J Radiol* 2004; 77(916):312-314.
16. Ciatto S, Houssami N, Apruzzese A et al. Reader variability in reporting breast imaging according to BI-RADS assessment categories (the Florence experience). *Breast* 2006; 15(1):44-51.
17. Berg WA, D'Orsy CJ, Jackson VP, Basset LW, Beam CA, Lewis RS. Does training in the breast imaging reporting and data system (BI-RADS) improve biopsy recommendations or feature analysis agreement with experienced breast imagers at mammography? *Radiology* 2002; 224:871–880.
18. Ciatto S, Houssami N, Apruzzese A et al. Categorizing breast mammographic density: intra- and interobserver reproducibility of BI-RADS density categories. *Breast* 2005; 14(4):269-75.

Table 1 – Rater agreement in each BIRADS 4 category

BIRADS Category		Rater			κ^*
	1 - n (%)	2 - n (%)	3 - n (%)	Consensus	
4A	308 (39.7)	318 (41.0)	80 (10.3)	235 (30.3)	1) 0.54 2) 0.39
4B	322 (41.5)	325 (41.9)	419 (54.1)	338 (43.6)	3) 0.38
4C	145 (18.8)	132 (17.1)	176 (35.6)	161 (20.8)	4) 0.44
Complete disagreement =				41 (5.3)	

*1= Raters 1 versus 2; 2 = 1 versus 3; 3 = 2 versus 3; 4 = overall

Table 2 – List of mammographic findings in each BIRADS 4 category and their respective positive predictive values (PPV)

BIRADS Category (consensus)	BI-RADS Descriptor	Benign n (%)	Malignant n (%)	PPV
4A	Calcification	109 (60.9)	36 (66.7)	24.8
	Mass	38 (20.9)	8 (14.8)	17.4
	Assimetry	23 (12.7)	4 (7.4)	14.8
	Architectural distortion	0	2 (3.7)	100
	Mass + Calcifications	5 (2.7)	2 (3.7)	28.6
	Assimetry + Calcifications	6 (3.3)	1 (1.8)	14.2
	Architectural distortion + Calcifications	0	1 (1.8)	100
	TOTAL 4A	181	54	23.0
4B	Calcification	93 (52.5)	75 (46.6)	44.6
	Mass	24 (13.6)	30 (18.6)	55.5
	Assimetry	35 (19.8)	27 (16.8)	43.5
	Architectural distortion	10 (5.6)	9 (5.6)	47.4
	Mass + Calcifications	6 (3.4)	9 (5.6)	60.0
	Assimetry + Calcifications	6 (3.4)	8 (4.9)	57.1
	Architectural distortion + Calcifications	3(1.7)	3 (1.8)	50.0
	TOTAL 4B	177	161	47.6
4C	Calcification	20 (45.4)	49 (41.9)	71.0
	Mass	1 (2.3)	10 (8.5)	90.1
	Assimetry	2 (4.6)	4 (3.4)	66.6
	Architectural distortion	15 (34.1)	33 (28.2)	68.7
	Mass + Calcifications	1 (2.3)	9 (7.7)	90.0
	Assimetry + Calcifications	1 (2.3)	7 (6.0)	87.5
	Architectural distortion + Calcifications	4 (9.1)	5 (4.2)	71.4
	TOTAL 4C	44	117	72.7
Complete disagreement	Calcification	15 (60)	10 (62.5)	40.0
	Mass	2 (8)	3 (18.7)	60.0
	Assimetry	3 (12)	0	0
	Architectural distortion	2 (8)	0	0
	Mass + Calcifications	0	2 (12.5)	100
	Assimetry + Calcifications	2 (8)	1 (6.2)	33.3
	Architectural distortion + Calcifications	1 (4)	0	0
	TOTAL compl. dis.	25	16	39.0
TOTAL (all categ.)		427	348	44.9

Table 3 – Interobserver variability in assigning the final BI-RADS 4 assessment category

<i>BIRADS Category</i>	<i>κ^*</i>
4A	0.27
4B	0.09
4C	0.39

Legends to figures:

1. Mammography shows amorphous calcifications enlarged on this true lateral view – consensus category 4B – Final pathological assessment: invasive ductal carcinoma.
2. Mammography demonstrates an oval, microlobulated mass with amorphous calcifications – consensus category 4B – Final pathological assessment: hyalinized fibroadenoma.
3. Mammography craniocaudal and lateral view of both breasts shows focal asymmetry at the superior upper quadrant of the left breast – consensus category 4B - Final pathological assessment: invasive lobular carcinoma.

Figure 1

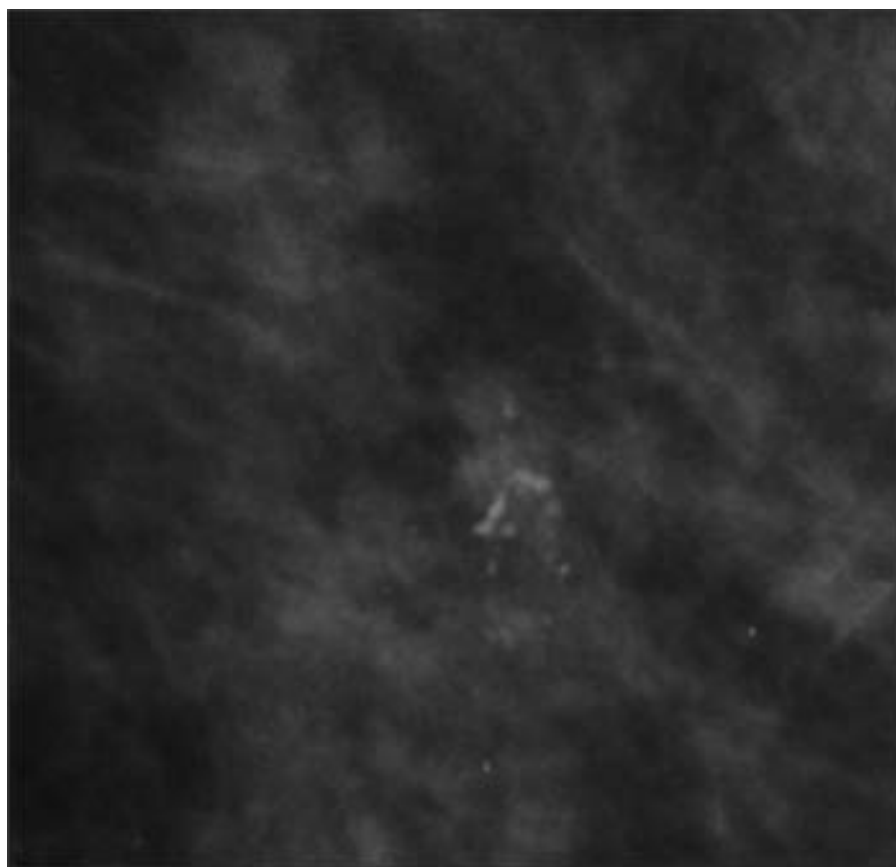


Figure 2

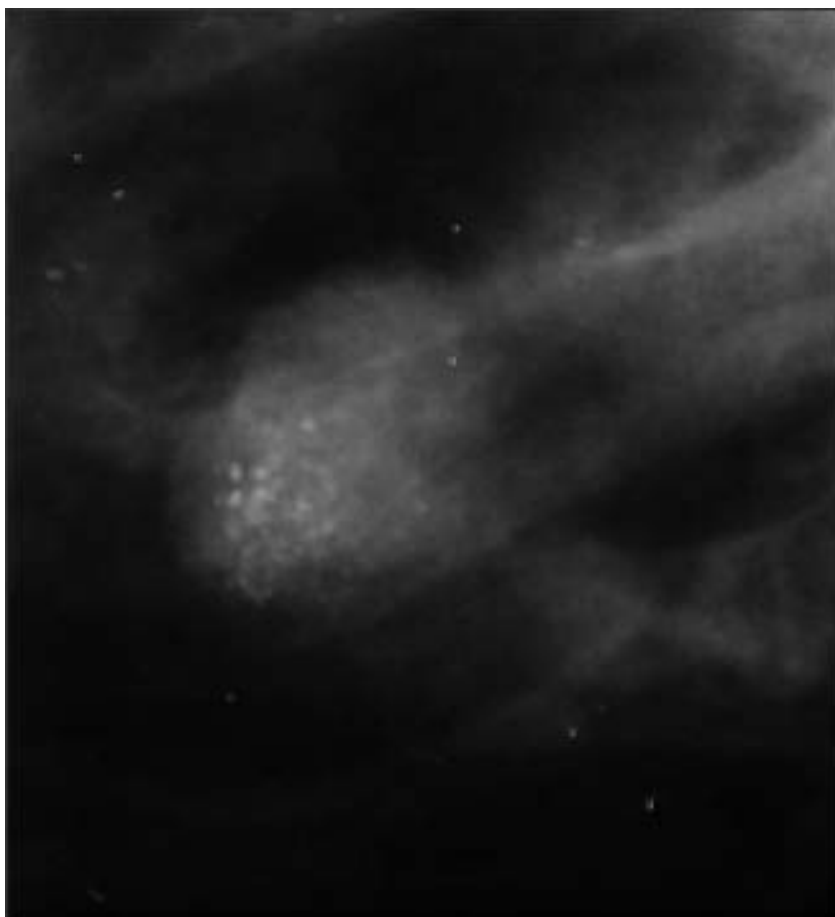
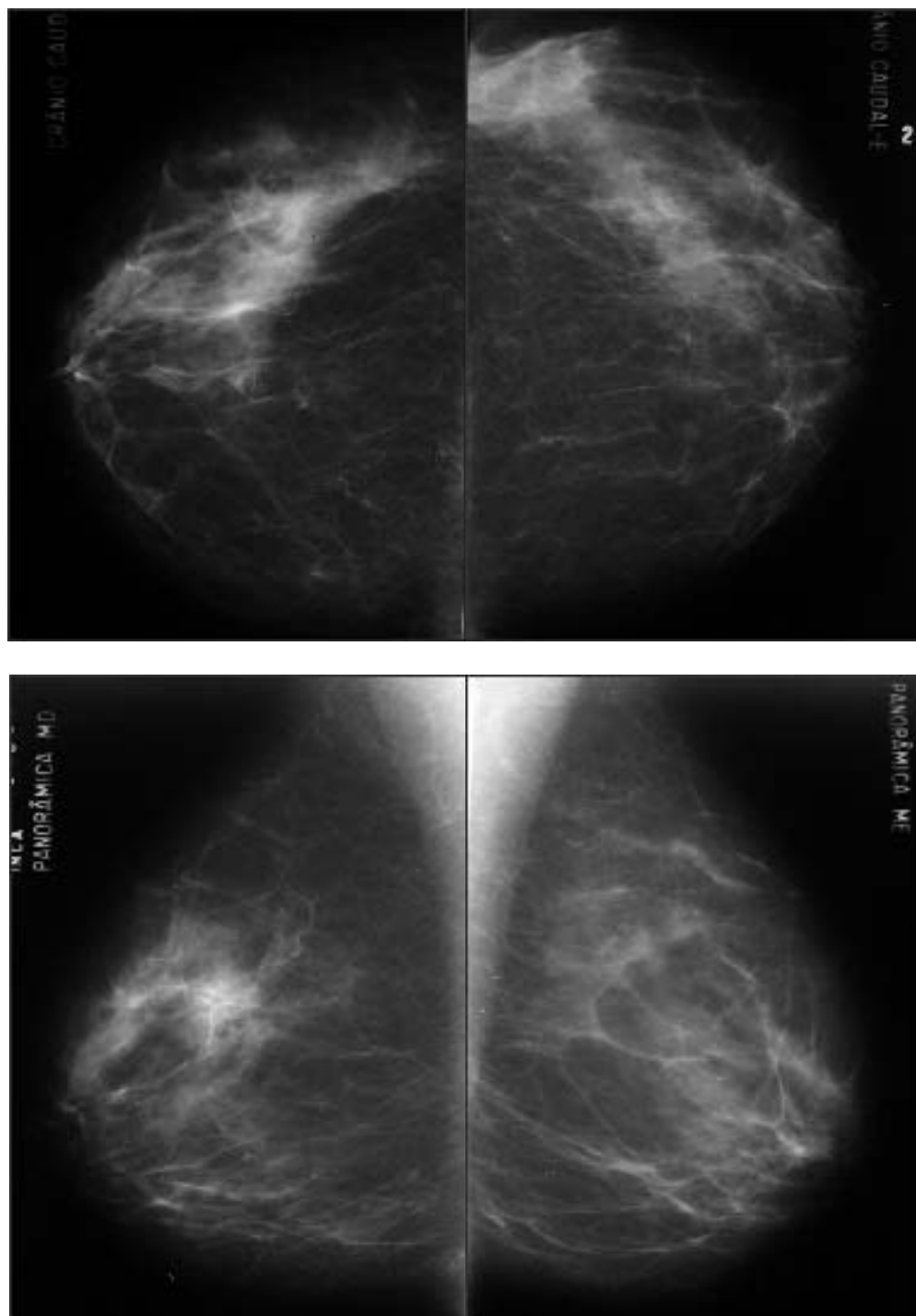


Figure 3



4. Conclusões

- O VPP das categorias 4A, 4B e 4C foi, respectivamente, de 23,0%, 47,6% e 72,7%.
- Houve moderada concordância geral entre os três observadores (0,4). A concordância foi pobre para a categoria 4A ($\kappa = 0,27$) e 4C ($\kappa = 0,39$) e mínima para a categoria 4B ($\kappa = 0,09$).

5. Referências Bibliográficas

American College of Radiology. Breast imaging report and data system (BI-RADS), Reston, VA: American College of Radiology; 1993.

American College of Radiology. Breast imaging report and data system (BI-RADS). 2nd ed. Reston, VA: American College of Radiology; 1995.

American College of Radiology. Breast imaging report and data system (BI-RADSTM). 3rd ed. Reston, VA: American College of Radiology; 1998.

American College of Radiology. Breast imaging report and data system (BI-RADS®). 4th ed. Reston, VA: American College of Radiology; 2003.

Ball CG, Butchart M, Macfarlane JK. Effect on biopsy technique of the breast imaging reporting and data system (BI-RADS) for nonpalpable mammographic abnormalities. Can J Surg 2002; 45(4):259-263.

Berg WA, Campassi C, Langenberg P, Sexton MJ. Breast Imaging Reporting and Data System: inter- and intraobserver variability in feature analysis and final assessment. AJR 2000; 174(6):1769-1777.

Bérubé M, Curpen B, Ugolini P, Lalonde L, Ouimet-Oliva D. Level of suspicion of a mammographic lesion: use of features defined by BI-RADS lexicon and correlation with large-core breast biopsy. Can Assoc Radiol J 1998; 49(4):223-228.

Brasil, Ministério da Saúde. Instituto Nacional de Câncer. Estimativa 2006 – Incidência de câncer no Brasil – 2006 - INCA [on line]. Rio de Janeiro. http://www.inca.gov.br/estimativa/2006/index.asp?link=conteudo_view.asp&ID=5 [acesso em 11 de outubro de 2007].

BRASIL. Ministério da Saúde. Conselho Nacional de Saúde – Resolução nº 196/96 sobre pesquisa envolvendo seres humanos. *Bioética*, 4(2): 15-25, 1996 (suplemento).

Ciatto S, Cataliotti L, Distante V. Non palpable lesions detected with mammography: review of 512 consecutive cases. *Radiology* 1987; 165:19-25.

Ciatto S, Houssami N, Apruzzese A, Bassetti E, Brancato B, Carozzi F et al. Reader variability in reporting breast imaging according to BI-RADS assessment categories (the Florence experience). *Breast* 2006; 15(1):44-51.

Cyrlak D. Induced costs of low-cost screening mammography. *Radiology* 1988; 168:661-663.

Hall FM, Storella JM, Siverstone DZ, Wyshak G. Non palpable breast lesions: recommendation for biopsy based on suspicion of carcinoma at mammography. *Radiology* 1988; 167:353-358.

Heywang-Köbruner SH, Dershaw DD, Scheer I. Diagnostic breast imaging: Mammography, sonography, magnetic resonance imaging, and interventional procedures. 2nd ed. New York: Ed Georg Thieme Verlag; 2001. P. 393.

Knutzen AM, Gisvold JJ. Likelihood of malignant disease for various categories of mammographically detected, nonpalpable breast lesions. *Mayo Clin Proc* 1993; 68(5):454-460.

Lacquement MA, Mitchell D, Hollingsworth AB. Positive predictive value of the Breast Imaging Reporting and Data System. J Am Coll Surg 1999; 189(1):34-40.

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33(1):159 -174.

Lazarus E, Mainiero MB, Schepps B, Koelliker SL, Livingston LS. BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value. Radiology 2006; 239(2):385-391.

Liberman L, Abramson AF, Squires FB, Glassman JR, Morris EA, Dershaw DD. The breast imaging reporting and data system: positive predictive value of mammographic features and final assessment categories. Am J Roentgenol 1998; 171(1):35-40.

Liberman L, Menell JH. Breast imaging reporting and data system (BI-RADS). Radiol Clin North Am 2002;40(3):409-430.

Margolin FR, Leung JW, Jacobs RP, Denny SR. Percutaneous imaging-guided core breast biopsy: 5 years' experience in a community hospital. Am J Roentgenol 2001; 177(3):559-564.

Mendez A, Cabanillas F, Echenique M, Malekshamran K, Perez I, Ramos E. Mammographic features and correlation with biopsy findings using 11-gauge stereotactic vacuum-assisted breast biopsy (SVABB). Ann Oncol 2004; 15(3):450-454.

Newman, L.A.; Sabel, M. Advances in breast cancer detection and management. Med Clin North Am 2003; 87(5):997-1028.

Orel SG, Kay N, Reynolds C, Sullivan DC. BI-RADS categorization as a predictor of malignancy. Radiology 1999; 211(3):845-850.

Smith-Bindman R, Chu PW, Miglioretti DL, Sickles EA, Blanks R, Ballard-Barbash R et al. Comparison of screening mammography in the United States and the United kingdom. JAMA 2003; 290(16):2129-2137.

Tan YY, Wee SB, Tan MP, Chong BK. Positive predictive value of BI-RADS categorization in an Asian population. Asian J Surg 2004; 27(3):186-191.

Tate PS, Rogers EL, Mcgee EM, Page GV, Hopkins SF, Shearer RG et al. Stereotactic breast biopsy: a six-year surgical experience. J Ky Med Assoc 2001; 99(3):98-103.

Thuler LC. Considerações sobre a prevenção do câncer de mama feminino. Rev. bras. Cancerol 2003; 49(4): 227-238.

Travade A, Isnard A, Bagard C, Bouchet F, Chouzet S, Gaillot A et al. Macrobiopsies stéréotaxiques par système à aspiration 11-G: à propos de 249 patientes. J Radiol 2002; 83(9 Pt 1):1063-1071.

World Medical Association – Declaration of Helsinki (1964), 2000. Disponível em: URL: <http://www.wma.net/e/policy/b3.htm>. Acessado em 02/12/2004.

Zonderland HM, Pope TL Jr, Nieborg AJ. The positive predictive value of the breast imaging reporting and data system (BI-RADS) as a method of quality assessment in breast imaging in a hospital population. Eur Radiol 2004; 14(10):1743-1750.

6. Anexos

6.1. Anexo 1 – Ficha para coleta de dados – Radiologista 1

Registro nº

Matrícula nº

BI-RADS categoria

6.2. Anexo 2 –Ficha para coleta de dados – Radiologista 2

Registro nº

Matrícula nº

BI-RADS categoria

6.3. Anexo 3 –Ficha para coleta de dados – Radiologista 3

Registro nº |__|__|__|__|

1) Achado mamográfico

- a) Nódulo: resente ☐ ausente ☐
- Forma: edonda ☐ oval ☐ lobulada ☐ irregular ☐
- Margem: circunscrita ☐ microlobulada ☐ obscurecida ☐
 indistinta ☐ espiculada ☐
- Densidade: alta densidade ☐ isodenso ☐ baixa densidade ☐
- b) Microcalcificações: presente ☐ ausente ☐
- Distribuição: difusas ☐ lineares ☐ agrupadas ☐ regionais ☐
 segmentares ☐
- Forma: puntiformes ☐ indistintas ou amorfas ☐ heterogêneas ☐
- c) Distorção da arquitetura: presente ☐ ausente ☐
- d) Assimetria: presente ☐ ausente ☐
 - focal ☐ difusa ☐
- e) Lesão associada: nódulo com microcalcificação ☐
 assimetria com microcalcificação ☐
 distorção arquitetural com microcalcificação ☐

2) Histopatológico: _____

6.4. Anexo 4 - Carta de aprovação do projeto CP – SPC – DPQ



COORDENAÇÃO DE PESQUISA
Serviço de Pesquisa Clínica - SPC/CPQ
Comissão Científica de Estudos Clínicos

Rio de Janeiro, 10 de fevereiro de 2005

Memo. 010/05 – CCEC

Da: CCEC

Para: Fabíola Procacci Kestelman

Senhor (a) Pesquisador (a):

Informamos que a Comissão Científica de Estudos Clínicos aprovou em sua reunião de 04/02/2005, o projeto de pesquisa nº 003/05 – Título: Associação entre os Critérios Morfológicos das Lesões não Palpáveis da Mama e a Subdivisão da Categoria 4 do BI-RADS, do qual o(a) Sr.(a) é o(a) Pesquisador(a) Principal.

Informamos que os projetos de pesquisa aprovados pela CCEC deverão ser objeto de: 1. Relatórios semestrais, a partir da data de assinatura do contrato com a FAF, ou no caso de estudos internos, a partir da data de aprovação pelo Comitê de Ética em Pesquisa. 2. Apresentações anuais em reuniões marcadas pela CPQ especificamente para este fim.

De acordo com o fluxograma vigente, o protocolo está sendo encaminhado ao Comitê de Ética em Pesquisa.

Atenciosamente,

Marija Maria Dreyer Breitenbach

CC – Dr. Luis Otávio Olivato
Presidente do Comitê de Ética em Pesquisa
Anexos para o CEP:
1 – Protocolo Clínico
2 – Folha de Rosto
3 – Curriculum Vitae
4 – Formulário da Pesquisa

6.5. Anexo 5 - Cartas de aprovação do projeto no CEP – INCA

PARECER DO COMITÊ DE ÉTICA EM PESQUISA

TÍTULO DO PROJETO: 03/05 - Associação entre os Critérios Morfológicos das Lesões não Palpáveis da Mama e a Subdivisão da Categoria 4 do BI-RADS	
PESQUISADORA RESPONSÁVEL: Fabíola Procaci Kestelman	
INSTITUIÇÃO: INSTITUTO NACIONAL DE CâNCER	DATA DE APRESENTAÇÃO AO CEP: _04_ / _02_ / _2005_
OBJETIVO: Objetivo Geral: Determinar o VPP dos critérios morfológicos em mamografias de mulheres com lesões não palpáveis e verificar sua associação com as categorias 4A, 4B e 4C propostas na 4ª edição do BI-RADS®, tendo como padrão ouro o estudo histopatológico.	
PARECER DO CEP: O Comitê de Ética decidiu por aprovar o protocolo acima sem restrições, não havendo a necessidade de obtenção do Termo de Consentimento Livre e Esclarecido	
MEMBROS DO CEP / INCA QUE PARTICIPARAM DO JULGAMENTO:	
Adriana Alves da Souza Scheliga Adriana Bonomo André Marcelo Machado Soares Carlos Frederico de Freitas Lima Claudio Calazani do Carmo	Dinah Schuman Edina Bajona Luis Cláudio Oliveira Marcelo Ribeiro Schimela Vânia Maria Fernandes Taborin
DATA DA REUNIÃO _28_ / _02_ / _2005_	ASSINATURA DO PRESIDENTE DO CEP: 



MINISTÉRIO DA SAÚDE
Comitê de Ética em Pesquisa-CEP

Rio de Janeiro, 15 de março de 2003

Dra. Fabíola Procaci Kestelman
Pesquisadora Principal

Ref. Prot. nº 03/05 – Associação entre os Critérios Morfológicos das Lesões não Palpáveis da Mama e a Subdivisão da Categoria 4 do BI-RADS

Prezada Doutora,

Informamos que o Comitê de Ética em Pesquisa do Instituto Nacional de Câncer após análise decidiu por aprovar o Protocolo intitulado: Associação entre os Critérios Morfológicos das Lesões não Palpáveis da Mama e a Subdivisão da Categoria 4 do BI-RADS em sua reunião de 28 de fevereiro de 2003.

Estamos encaminhando a documentação pertinente para o CONEP, com vistas a registro e arquivamento.

Atenciosamente,


Dr. Luis Otávio Olivatto
Coordenador do Comitê de Ética em Pesquisa
CEP-INCA

CC: Dra. Elyete de Oliveira Canella
Chefe do Serviço de Radiologia – H CIII



MINISTÉRIO DA SAÚDE
CONM de Ética em Pesquisa-INCA

Rio de Janeiro, 15 de março de 2005

Dr. William Saad Hoene
Coordenador do Grupo de Trabalho - CONEP-MS
Ministério da Saúde - Conselho Nacional de Saúde
Bloco G - Anexo - Ala B - 1º andar - Salas 128 e 147
70058-901 - Brasília - DF

Ref. Prot. nº 03/05 – Associação entre os Critérios Morfológicos das Lesões não Palpáveis da Mama e a Subdivisão da Categoria 4 do BI-RADS

Prezado Doutor,

De acordo com o Decreto nº 109, de 02 de maio de 1991, publicado no D.O.U. de 03 de maio de 1991, no seu artigo 14, § IV, ao Instituto Nacional de Câncer, INCA, compete coordenar, programar e realizar pesquisas clínicas, epidemiológicas e experimentais em cancerologia.

Para tanto o INCA possui uma ampla infra-estrutura com Laboratório de Análises Clínicas; Serviço de Imagens de Radiologia, Tomografia Computadorizada, Ressonância Magnética; tratamento multidisciplinar com os Serviços de Cirurgia, Radioterapia e Quimioterapia; Centro de Transplante de Medula Óssea; Serviço Terapêutico Oncológico (CSTO, para pacientes fora de possibilidades terapêuticas); Corpo clínico, cirúrgico e de enfermagem com ampla experiência em cancerologia; estrutura hospitalar adequada para internações com serviço de urgência e emergência e serviço de UTI.

O INCA é o principal Centro do Ministério da Saúde em cancerologia do Brasil.

Atenciosamente,



Dr. Luis Cláudio Olivatto
Coordenador do Comitê de Ética em Pesquisa
CER-INCA