



## Redes Bayesianas e Regressão Logística em pacientes submetidas a core biópsia

Bayesian Networks and Logistic Regression in patients undergoing core biopsy

Redes Bayesianas e Regressão Logística em pacientes sometidos a biópsia de núcleo

Gustavo de Agostin<sup>1</sup>, Ricardo Rampinelli Formigoni<sup>2</sup>, Larissa Mayumi Yokoi<sup>3</sup>, Erik Paul Winnikow<sup>4</sup>, Priscyla Waleska Simões<sup>5</sup>

### RESUMO

**Descritores:** Câncer de Mama; Sistemas de Apoio a Decisões Clínicas; Redes Bayesianas

**Objetivos:** Buscou-se correlacionar os achados dos exames de imagem com o resultado histopatológico pela elaboração de uma Rede Bayesiana. **Métodos:** Foi realizado um estudo do tipo transversal, envolvendo 164 pacientes do sexo feminino, totalizando 170 nódulos, submetidas a Core Biópsia. **Resultados:** A aquisição de conhecimento para a explicitação da parte quantitativa e qualitativa da Rede Bayesiana foi realizada por meio do aprendizado bayesiano que ocorreu por meio do algoritmo expectation maximization. A sensibilidade da Rede Bayesiana foi 93,3% (IC 95%: 82,9-96,0), a especificidade foi 88,6% (IC 95%: 81,1-96,0) e a acurácia foi 89,4%. Após a análise multivariada, os fatores de risco associados ao câncer de mama foram: idade maior ou igual a 50 anos e nódulo maior que dois centímetros. **Conclusão:** Mediante resultados apresentados na análise multivariada, temos evidências que o conhecimento modelo pela Rede Bayesiana mostra-se coerente e indicativo de uso como apoio a decisão na prática clínica.

### ABSTRACT

**Keywords:** Breast Cancer; Decision Support Systems; Clinical; Bayesian Networks

**Objectives:** It was sought to correlate the findings of imaging with histopathological for the elaboration of a Bayesian Network. **Methods:** A cross-sectional study was involving 164 female patients, totaling 170 nodules, submitted to Core Biopsy. **Results:** The knowledge acquisition for the explanation of the quantitative and qualitative part of Bayesian Network was accomplished through Bayesian learning that occurred through the expectation maximization algorithm. The sensitivity of Bayesian Network was 93.3% (CI 95%: 82.9 to 96.0), the specificity was 88.6% (CI 95%: 81.1 to 96.0) and the accuracy was 89.4%. After multivariate analysis, risk factors associated with breast cancer were age greater than or equal to 50 years and nodule larger than two centimeters. **Conclusion:** Through results presented in the multivariate analysis, we have evidence that the model knowledge by the Bayesian Network is coherent and indicative of use as a support for decision in clinical practice.

### RESUMEN

**Descriptores:** Câncer de mama; Sistemas de Apoio a Decisões Clínicas; Redes bayesianas

**Objetivos:** Él trató de correlacionar los hallazgos de los estudios de imagen con resultados histopatológicos pela elaboração de uma Rede Bayesiana. **Métodos:** Un estudio transversal se llevó a cabo con la participación de 164 pacientes de sexo femenino, con un total de 170 nodos, sujeto a Core biópsia. **Resultados:** La adquisición de conocimiento para explicar la parte cuantitativa y cualitativa de la Red Bayesiana se logró a través del aprendizaje bayesiano que ocurrió a través del algoritmo de maximización de expectativas. La sensibilidad de la Red Bayesiana fue del 93,3% (IC del 95%: 82,9-96,0), la especificidad fue del 88,6% (IC del 95%: 81,1-96,0) y la precisión fue de 89,4%. **Conclusión:** Tras el análisis multivariado, los factores de riesgo asociados con el cáncer de mama fueron la edad mayor o igual a 50 años y superior a tanto alzado de dos centímetros. **Conclusión:** Con base en los resultados presentados en el análisis multivariado, tenemos evidencia de que el conocimiento del modelo por la Red Bayesiana es coherente e indicativo de uso como soporte para la decisión en la práctica clínica.

<sup>1</sup> Médico Residente do Programa de Residência Médica em Oncologia Clínica, Hospital de Clínicas de Porto Alegre, Porto Alegre (RS), Brasil.

<sup>2</sup> Médico Graduado do Curso de Medicina da Universidade do Extremo Sul Catarinense - UNESC, Criciúma (SC), Brasil.

<sup>3</sup> Mestranda do Programa de Pós-Graduação em Engenharia Biomédica, Centro de Engenharia, Modelagem e Ciências Sociais Aplicadas - CECS, Universidade Federal do ABC - UFABC, São Bernardo do Campo (SP), Brasil.

<sup>4</sup> Médico e Professor do Curso de Medicina da Universidade do Extremo Sul Catarinense - UNESC, Criciúma (SC), Brasil.

<sup>5</sup> Professora Adjunta do Programa de Pós-Graduação em Engenharia Biomédica, Centro de Engenharia, Modelagem e Ciências Sociais Aplicadas - CECS, Universidade Federal do ABC - UFABC, São Bernardo do Campo (SP), Brasil.

## INTRODUÇÃO

O câncer de mama é um problema de saúde global, já que cerca de um milhão de mulheres no mundo recebem o diagnóstico a cada ano, e mais de 410.000 mortes ocorrem devido a esta doença, representando 14% das mortes por câncer no sexo feminino<sup>(1)</sup>.

Aproximadamente 50% dos novos casos diagnosticados de câncer de mama podem ser explicados por fatores de risco conhecidos, como por exemplo, a idade da menarca, primeira gestação, menopausa e doenças proliferativas da mama. Em torno de 10% dos casos há associação com história familiar positiva. Ainda, o risco pode ser modificado por fatores sociodemográficos, ambientais e estilo de vida; porém essas associações ainda não foram claramente evidenciadas<sup>(2)</sup>.

O melhor método para rastreamento mamário consiste na realização da mamografia, pois permite diagnosticar a doença ainda na fase assintomática. A partir de 2002, a United States Preventive Services Task Force, baseado em estudos prévios, recomenda a realização da mamografia, com ou sem o exame clínico das mamas, a cada um ou dois anos para mulheres com 40 anos ou mais. O exame possui sensibilidade entre 77%-95% e a especificidade entre 94%-97%, sendo aplicável na maioria das mulheres<sup>(3)</sup>.

Nesse sentido, o diagnóstico patológico é fundamental no manejo das lesões mamárias. Atualmente a core biópsia substituiu a punção aspirativa por agulha fina para lesões palpáveis e para aquelas detectadas pelos exames de rastreamento, se tornou o exame padrão-ouro no diagnóstico pré-operatório do câncer de mama<sup>(4)</sup>.

Por outro lado, as Redes Bayesianas podem fornecer informações importantes que auxiliam o diagnóstico e prognóstico do câncer de mama. Por meio dessa análise, é possível representar as interações existentes entre as variáveis de interesse e medir quantitativamente o impacto sobre o desfecho (câncer de mama)<sup>(5)</sup>.

O objetivo desse trabalho foi correlacionar os achados dos exames de imagem com o resultado histopatológico pela elaboração de uma Rede Bayesiana.

## MÉTODOS

Foi realizado um estudo do tipo transversal, em um serviço de radiologia localizado no Sul do Brasil, envolvendo 164 mulheres, totalizando 170 nódulos, submetidas a Core Biópsia no período entre julho de 2014 e fevereiro de 2015, com pelo menos um exame de imagem prévio com classificação BI-RADS<sup>®</sup> 3, 4 ou 5, que aceitaram participar do estudo por meio de assinatura do termo de

consentimento. Todos os exames ultrassonográficos e coleta de material pela Core Biópsia foram realizados pelo mesmo profissional. O estudo foi aprovado pelo Comitê de Ética do local de origem da pesquisa, parecer 700.772.

Anteriormente à realização do exame de imagem, foi preenchido um instrumento com as variáveis estudadas: idade, história familiar de primeiro grau para câncer de mama, palpabilidade do nódulo ao exame físico, tamanho do nódulo pela ultrassonografia, classificação BI-RADS<sup>®</sup> na mamografia e ultrassonografia, e o resultado histopatológico. A coleta de dados foi realizada por um profissional treinado previamente.

Os dados foram analisados com auxílio do software IBM Statistical Package for the Social Sciences (SPSS) versão 26.0. As variáveis quantitativas foram expressas como média e desvio padrão ou dicotomizadas e expressas por meio de frequência absoluta e relativa. As variáveis qualitativas foram expressas por meio de frequência absoluta e relativa.

As análises inferenciais foram realizadas com um nível de significância  $\alpha = 0,05$  e intervalo de confiança de 95%. A investigação da existência de associação entre as variáveis preditoras e o desfecho foi realizada por meio da aplicação do Teste qui-quadrado de Pearson e Teste Exato de Fisher. A medida da força de associação calculada foi o Odds Ratio (OR) e realizada a Regressão Logística Binária.

A Rede Bayesiana (RB) foi modelada na ferramenta Netica versão 6.07 (<http://www.norsys.com>), por ser uma ferramenta intuitiva, que apresenta bons resultados em estudos com Redes Bayesianas<sup>(6)</sup> sendo tal estrutura composta pelo perfil da paciente, sinais e exames relacionados ao câncer de mama. A aquisição de conhecimento para a explicação da parte quantitativa e qualitativa da RB foi realizada por meio do aprendizado bayesiano que ocorreu por meio do algoritmo expectation maximization (EM). Um problema que as RB se propõem a resolver é a falta de informações, algo que acontece na maioria dos datasets. Nesse contexto, um dos métodos mais utilizados para aprendizagem da Rede Bayesiana é o algoritmo EM<sup>(7)</sup>.

Para elaboração da RB, foram realizados experimentos com o dataset dividido na proporção de 2/3 para o treinamento e 1/3 para testes, no entanto o modelo foi melhor gerado dividido em um conjunto de treinamento (n=85 nódulos, 23 malignas) e um conjunto de testes (n=85 nódulos, 22 malignas). Após a etapa de treinamento, os atributos foram definidos como nós da RB, conforme apresenta a Tabela 1.

Ao final, foi realizada a validação da Rede Bayesiana, sendo calculada sensibilidade, especificidade, acurácia e respectivos intervalos de confiança.

**Tabela 1** – Descrição e Estados dos Nós

Descrição dos Nós	Estados dos Nós
Desfecho	Maligno, Benigno
História Familiar	Sim, Não
Nódulo Palpável	Sim, Não
BI-RADS <sup>®</sup> USG	BI-RADS <sup>®</sup> 3, BI-RADS <sup>®</sup> 4a, BI-RADS <sup>®</sup> 4b, BI-RADS <sup>®</sup> 4c, BI-RADS <sup>®</sup> 5
BI-RADS <sup>®</sup> Mamografia	BI-RADS <sup>®</sup> 0, BI-RADS <sup>®</sup> 1, BI-RADS <sup>®</sup> 2, BI-RADS <sup>®</sup> 3, BI-RADS <sup>®</sup> 4, BI-RADS <sup>®</sup> 5
Idade	Maior ou igual a 50 anos, Menor que 50 anos
Tamanho do nódulo	Maior que 2 cm, Menor ou igual a 2 cm

BI-RADS<sup>®</sup> = Breast Imaging-Reporting and Data System

## RESULTADOS E DISCUSSÃO

A amostra total estudada foi composta por 170 nódulos, pertencentes a 164 mulheres com média de idade de 46,13 ( $\pm 14,19$ ) anos, sendo a idade mínima de 19 anos e a máxima de 89 anos. A Tabela 2 apresenta as características gerais da amostra.

**Tabela 2** – Características gerais da amostra

Variável	n (%)
Idade	
≥ 50 anos	63 (37,1)
< 50 anos	107 (62,9)
Histórico Familiar	
Sim	16 (9,4)
Não	154 (90,6)
Nódulo Palpável	
Sim	77 (45,3)
Não	93 (54,7)
Tamanho do nódulo	
> 2 cm	56 (33,0)
≤ 2 cm	114 (67,0)
BI-RADS® Mamografia	
0	28 (16,47)
1	27 (15,88)
2	6 (3,53)
3	20 (11,76)
4	27 (15,88)
5	18 (10,59)
NR	44 (25,88)
BI-RADS® Ultrassonografia	
3	50 (29,41)
4a	68 (40,00)
4b	13 (7,65)
4c	8 (4,70)
5	31 (18,23)
Idade Média (DP)	46,13 ( $\pm 14,19$ )
Tamanho médio do nódulo (DP)	1,92 ( $\pm 1,39$ )

BI-RADS® = Breast Imaging-Reporting and Data System; DP = Desvio Padrão

Em relação à idade, quando analisadas as pacientes com 50 anos ou mais, a chance de malignidade do nódulo foi 3,71 (IC 95%: 1,82-7,55) vezes maior quando comparadas às menores de 50 anos. Foi visto também que mulheres com nódulo palpável (ao exame físico) ou maior que dois centímetros (pela ultrassonografia) apresentaram chance 9,83 (IC 95%: 4,19-23,03) e 3,78 (IC

95%: 1,85-7,76) vezes maior de malignidade, respectivamente (Tabela 3).

A Tabela 4 apresenta o modelo final resultante da análise multivariada, que revelou aumento do risco para a idade e tamanho do nódulo. Assim, mulheres com idade maior ou igual a 50 anos (OR = 5,47; IC 95%: 2,31-12,96) e nódulo maior que dois centímetros (OR = 12,99; IC 95%: 5,08-33,22) tiveram maior associação com o câncer de mama.

A partir das informações obtidas no dataset foi elaborada a RB cuja apresentação geral pode ser observada na Figura 1A.

Mediante resultados apresentados na análise multivariada, temos evidências que o conhecimento modelado pela Rede Bayesiana mostra-se coerente e indicativo de uso como apoio à decisão na prática clínica, ou seja, ao selecionar os fatores de risco elucidados na análise multivariada (idade maior ou igual a 50 anos e nódulo maior que dois centímetros), pode-se observar que a probabilidade de nódulo maligno passa de 26,7% para 59,9%, conforme ilustra a figura 1-B.

Em um aspecto geral, sem a propagação de evidências, a probabilidade de malignidade do nódulo é de 26,7% (Figura 1-A). Considerando as seguintes variáveis: história familiar negativa, nódulo palpável ao exame físico, idade maior ou igual a 50 anos, nódulo maior que dois centímetros, BI-RADS® 0 na mamografia e BI-RADS® 4B na ultrassonografia, a RB mostrou uma probabilidade de 56,4% de malignidade do nódulo (Figura 2-A). Ao selecionar as mesmas variáveis acima, apenas substituindo o valor do BI-RADS® da ultrassonografia para 4C, a probabilidade de o nódulo ser maligno aumenta para 97,8% (Figura 2-B).

Em contrapartida, ao selecionarmos as variáveis a seguir: história familiar negativa, nódulo não palpável ao exame físico, idade menor que 50 anos, nódulo menor ou igual a 2cm, BI-RADS® 0 na mamografia e BI-RADS® 4C na ultrassonografia, obteve-se uma probabilidade de 84,1% de benignidade do nódulo (Figura 2-C). Alternando o valor do BI-RADS® do ultrassom para 4B, a probabilidade de benignidade do nódulo subiu para 99,5% (Figura 2-D).

Mediante testes realizados, a sensibilidade da Rede Bayesiana foi 93,3% (IC 95%: 82,9-96,0), a especificidade

**Tabela 3** – Relação de Variáveis Independentes com o Desfecho

Variável	Nódulo n (%)			OR	IC 95%	p
	Total (n = 170)	Maligno (n = 45)	Benigno (n = 125)			
Idade						
≥ 50 anos	63 (37,1)	27 (60,0)	36 (28,8)			
< 50 anos	107 (62,9)	18 (40,0)	89 (71,2)	3,71	[1,82; 7,55]	<0,001
Histórico Familiar						
Sim	16 (9,4)	3 (6,7)	13 (10,4)	-	-	0,50
Não	154 (90,6)	42 (93,3)	112 (89,6)			
Nódulo Palpável						
Sim	77 (45,3)	37 (82,2)	40 (32,0)	9,83	[4,19; 23,03]	<0,001
Não	93 (54,7)	8 (17,8)	85 (68,0)			
Tamanho do nódulo						
> 2 cm	56 (33,0)	24 (53,3)	29 (23,2)	3,78	[1,85; 7,76]	<0,001
≤ 2 cm	114 (67,0)	21 (46,7)	96 (76,8)			

OR = Odds Ratio; IC= Intervalo de Confiança

Tabela 4 – Análise multivariada

Variável	OR <sub>ajustado</sub>	CI 95%	p
Idade			
≥ 50 anos	5,473	[2,312 – 12,959]	<0,001
< 50 anos	1		
Tamanho do nódulo			
> 2 cm	12,993	[5,082 – 33,220]	<0,001
≤ 2 cm	1		

OR = Odds Ratio; IC= Intervalo de Confiança

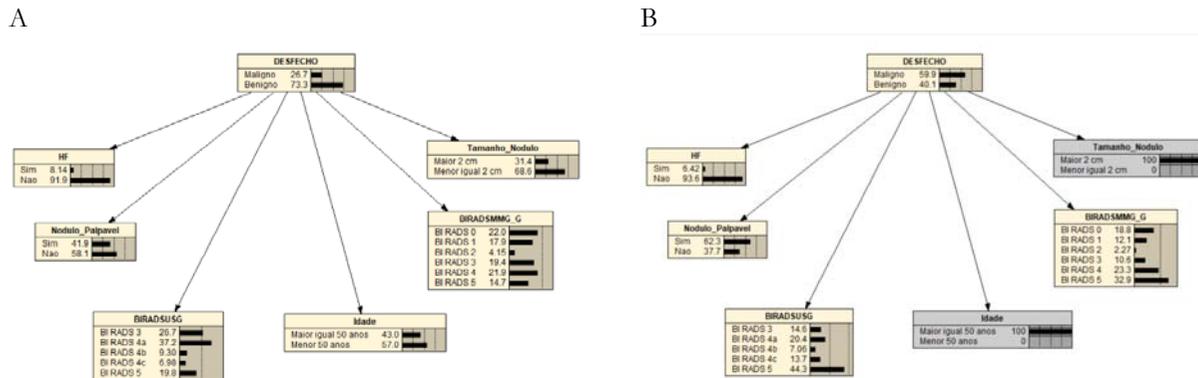


Figura 1 – Rede Bayesiana

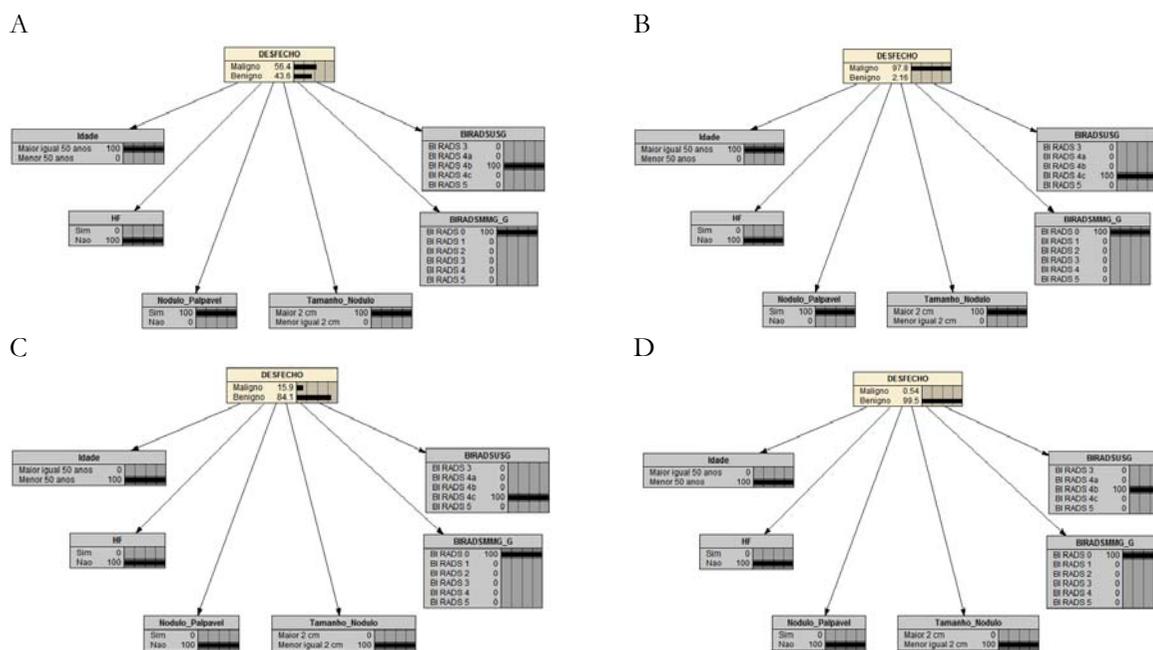


Figura 2 – Propagação de evidências na Rede Bayesiana da pesquisa

foi 88,6% (IC 95%: 81,1 - 96,0) e a acurácia foi 89,4%.

A mamografia é um importante método para a detecção precoce do câncer de mama, porém muitos achados não podem ser facilmente classificados como malignos ou benignos. O sucesso diagnóstico depende da habilidade do médico em detectar anormalidades mamográficas e integrar as informações clínicas aos fatores de risco e achados no exame físico para determinar a probabilidade do diagnóstico de câncer de mama<sup>(8)</sup> idade maior ou igual a 50 anos foi considerado fator de risco (OR=3,71; IC 95%: 1,82-7,55) para o câncer de mama. Em um estudo alemão de 2015, criou-se uma rede bayesiana que avaliou 1276 nódulos, essa rede demonstrou que 86,9% (n=120) dos nódulos malignos ocorreram em mulheres com 50 anos ou mais<sup>(9)</sup>. Em nosso trabalho 60% das neoplasias malignas estavam compreendidas nessa faixa

etária.

A idade da mulher representa um importante fator de risco para o câncer de mama. Tal fato pode ser explicado pela presença de estrogênio (e outros carcinógenos) que induz à mutação do DNA. Assim, com o avançar da idade, tais mutações tendem a se acumular e consequentemente aumentam o risco de desenvolver o câncer de mama<sup>(10)</sup>. A idade como fator de risco foi também evidenciada por relatórios publicados pela base de dados Surveillance, Epidemiology, and End Results (SEER) filiada ao National Cancer Institute (NCI) dos Estados Unidos. Tal revisão demonstrou que a probabilidade de uma mulher desenvolver câncer de mama entre 2006 e 2008 naquele país foi: 1 para cada 203 mulheres até 39 anos; 1 em 27 dos 40 aos 59 anos; 1 em 28 dos 60 aos 69 anos; e 1 em 15 nas mulheres com 70 anos ou mais<sup>(11)</sup>.

Em uma série prospectiva realizada na Alemanha, em 1975, avaliando 1244 pacientes sintomáticas, o exame físico (incluindo palpação do nódulo), mostrou-se eficaz em definir o câncer de mama em 93% dos casos<sup>(12)</sup>. No nosso trabalho, concluímos que pacientes com nódulo palpável ao exame físico, isoladamente, obtiveram chance 9,83 (IC: 4,19-23,03) vezes maior de câncer de mama. Esses dados mostram a importância de um exame físico bem realizado, que pode, mesmo antes dos exames complementares, nortear o raciocínio médico para um diagnóstico mais provável.

Considerando-se o tamanho do nódulo ao ultrassom, nossa casuística demonstrou que entre os nódulos malignos, 53,3% (n=24) eram maiores que 2cm, ao passo que, entre os nódulos benignos 76,8% (n=96) foram menores ou igual a 2 cm. Em estudo prospectivo realizado na China entre 1996 e 1998, avaliou 1203 pacientes com tumor palpável de mama. Tal estudo apresentou achados semelhantes ao nosso, sendo que entre os nódulos malignos 50,13% (n=196) eram maiores que 2cm e quando avaliados apenas os benignos, 62,44% (n=507) eram menores ou igual a 2cm<sup>(13)</sup>.

Em uma meta-análise britânica, de 2001, reunindo 52 estudos epidemiológicos, foram avaliadas 58.209 mulheres com câncer de mama e 101.986 sem a doença; concluiu-se que o risco para câncer de mama foi fortemente associado com o número de casos em parentes de primeiro grau, sendo o dobro do risco se a mulher possuir um caso diagnosticado entre um parente de primeiro grau e três vezes maior caso haja dois casos na família<sup>(14)</sup>. Nosso estudo mostrou que a história familiar de primeiro grau não apresentou aumento do risco de desenvolver a neoplasia.

Todos os nódulos mamários de nossa casuística com classificação BI-RADS<sup>®</sup> 2 ou 3, em um dos exames de imagem, foram benignos ao anatomopatológico, e todas as lesões classificadas como BI-RADS<sup>®</sup> 5 foram malignas. Dados semelhantes podem ser observados em um estudo sequencial e prospectivo realizado no Brasil em 2001, com 197 nódulos avaliados, que revelou benignidade em 100% (n=5) e 98,8% (n=85) dos casos classificados como BI-RADS<sup>®</sup> 2 e 3, respectivamente, e obteve malignidade de 97,2% (n=35) das lesões quando classificadas como estágio 5 de BI-RADS<sup>®</sup><sup>(15)</sup>. O estudo apresentado na Sociedade Europeia de Radiologia<sup>(9)</sup>, citado anteriormente, também sugeriu resultados similares com 100% de benignidade para nódulos BI-RADS<sup>®</sup> 2 (n=431) e BI-RADS<sup>®</sup> 3 (n=77) e 93,85% de malignidade (n=61) daqueles classificados como BI-RADS<sup>®</sup> 5 no exame mamográfico.

Segundo a American College of Radiology, a classificação BI-RADS<sup>®</sup> 4 no exame de imagem denota uma anormalidade suspeita e apresenta um risco de malignidade entre 2% e 94%<sup>(16)</sup>. Por ser uma categoria muito ampla, subdivisões foram criadas para um manejo mais adequado, sendo essas 4A (probabilidade de malignidade de 2%-9%), 4B (probabilidade de malignidade 10%-49%) e 4C (probabilidade de malignidade de 50%-94%)<sup>(16)</sup>. Em nosso estudo, ao exame ultrassonográfico das mamas, os nódulos classificados como BI-RADS<sup>®</sup> 4A, 4B e 4C apresentaram-se malignos ao anatomopatológico em 5,9%, 30,8% e 75%, respectivamente.

O Ministério da Saúde orienta a iniciar o rastreamento do câncer de mama a partir dos 50 anos de idade, não sendo indicada a mamografia como rastreamento para mulheres entre 40 e 49 anos<sup>(17)</sup>. Entre os nódulos malignos de nossa amostra (n=45), 18 (40%) apresentaram-se em mulheres com idade inferior a 50 anos, sendo que 12 (26,66%) desses nódulos foram encontrados dos 40 aos 49 anos. Em uma meta-análise que avaliou mulheres entre 40 e 49 anos, submetidas à mamografia como método diagnóstico por um período médio de 14 anos de seguimento, notou-se 15% de redução na mortalidade por câncer de mama (RR=0,85; IC<sub>95%</sub>: 0,73-0,99)<sup>(18)</sup>. Os dados supracitados não condizem com as orientações adotadas no Brasil atualmente, pois demonstram a necessidade em iniciar o rastreamento do câncer de mama em idade mais precoce. No entanto, são necessários mais estudos e discussões a respeito do tema já que o adiamento do diagnóstico pode ser prejudicial a muitas pacientes.

As Redes Bayesianas permitem identificar visualmente quais atributos podem representar uma influência direta sobre o resultado e como tais atributos estão relacionados entre si<sup>(19)</sup>. O propósito de uma Rede Bayesiana é aumentar, e não substituir, os recursos diagnósticos no complexo processo de diagnóstico clínico<sup>(20)</sup>.

O uso do algoritmo EM também pode ser observado em outros estudos da área médica, como na pesquisa Brasileira<sup>(21)</sup>, realizada em 2012, que avaliou o número de óbitos ocorridos em 2000 naquele local a partir de dados disponibilizados via Sistema de Informação sobre Mortalidade (SIM), sendo utilizadas duas metodologias, o estimador bayesiano empírico e o algoritmo EM, a fim de estimar um fator de correção que, aplicados aos óbitos registrados, corrigisse as subnotificações e fosse gerado um número de óbitos viável para a elaboração de indicadores de mortalidade confiáveis. Para o total do Estado avaliado, foi estimado um sub-registro de 11% com o estimador bayesiano empírico e 12,9% quando utilizado o algoritmo EM<sup>(21)</sup>.

Há duas décadas surgiram os primeiros estudos utilizando RB e aplicados ao câncer de mama. Um dos estudos pioneiros foi elaborado por Kahn et al, em 1995<sup>(22)</sup>. Esse estudo americano, denominado MammoNet, utilizou cinco achados de anamnese, dois achados de exame físico e 15 achados mamográficos na elaboração da RB como ferramenta auxiliar na decisão diagnóstica. Outro estudo americano precursor, publicado em 2004, fez uso de uma RB para avaliar a probabilidade de malignidade de microcalcificações mamárias detectadas na mamografia e sua capacidade de invasão através de características clínicas e mamográficas das pacientes<sup>(23)</sup>. Nos últimos anos, diversos estudos já modelaram RB no contexto do câncer de mama.

O resultados de sensibilidade (93,3%), especificidade (88,6%) e AUC (0,894) obtidos com a nossa Rede Bayesiana aproximam-se aos resultados de estudos com características similares apresentados anteriormente<sup>(22, 24-25)</sup>.

## CONCLUSÃO

Nosso estudo mostrou que todos os fatores em estudo

foram associados ao Câncer de Mama, com exceção do histórico familiar; e maiores nódulos em mulheres com mais de 50 anos foram os fatores de risco conjuntamente associados ao Câncer de Mama.

Mediante resultados apresentados na análise multivariada, temos evidências que o conhecimento modelado pela Rede Bayesiana se mostra coerente e indicativo de uso como apoio a decisão na prática clínica.

No entanto, sugerimos a inclusão de outras variáveis em futuros modelos bayesianos, como densidade das mamas, formato das lesões e história de reposição

hormonal, que poderá melhorar a performance do modelo apresentado nesse estudo.

## AGRADECIMENTOS

Esse estudo foi financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Universidade Federal do ABC (UFABC) e Universidade do Extremo Sul Catarinense (UNESC).

## REFERÊNCIAS

- Coughlin SS, Ekwueme DU. Breast cancer as a global health concern. *Cancer epidemiology*. 2009;33(5):315-8.
- Factors that modify breast cancer risk in women [Internet]. UpToDate. 2020 [cited 20/09/2020]. Available from: <http://www.uptodate.com/contents/factors-that-modify-breast-cancer-risk-in-women>.
- U.S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Annals of internal medicine*. 2009;151(10):727-37.
- Chen X, Sun L, Mao Y, Zhu S, Wu J, Huang O, et al. Preoperative core needle biopsy is accurate in determining molecular subtypes in invasive breast cancer. *BMC cancer*. 2013;13:390.
- Cruz-Ramírez N, Acosta-Mesa HG, Carrillo-Calvet H, Nava-Fernández LA, Barrientos-Martínez RE. Diagnosis of breast cancer using Bayesian networks: a case study. *Computers in biology and medicine*. 2007;37(11):1553-64.
- Fuster-Parra P, Vidal-Conti J, Borràs PA, Palou P. Bayesian networks to identify statistical dependencies. A case study of Spanish university students' habits. *Informatics for health & social care*. 2017;42(2):166-79.
- Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1977;39(1):1-22.
- Feig SA. Breast masses: Mammographic and sonographic evaluation. *Radiologic clinics of North America*. 1992;30(1):67-92.
- Benndorf M, Kotter E, Langer M, Herda C, Wu Y, Burnside ES. Development of an online, publicly accessible naive Bayesian decision support tool for mammographic mass lesions based on the American College of Radiology (ACR) BI-RADS lexicon. *European radiology*. 2015;25(6):1768-75.
- Canada. Review of Lifestyle and Environmental Risk Factors for Breast Cancer: Summary Report. Canadian Breast Cancer Initiative. Working Group on Primary Prevention of Breast Cancer: Depository Services Program; 2001.
- Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA: a cancer journal for clinicians*. 2012;62(1):10-29.
- Rimsten A, Stenkvist B, Johanson H, Lindgren A. The diagnostic accuracy of palpation and fine-needle biopsy and an evaluation of their combined use in the diagnosis of breast lesions: report on a prospective study in 1244 women with symptoms. *Annals of surgery*. 1975;182(1):1-8.
- Chen SC, Cheung YC, Su CH, Chen MF, Hwang TL, Hsueh S. Analysis of sonographic features for the differentiation of benign and malignant breast tumors of different sizes. *Ultrasound in obstetrics & gynecology: the official journal of the International Society of Ultrasound in Obstetrics and Gynecology*. 2004;23(2):188-93.
- Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet (London, England)*. 2001;358(9291):1389-99.
- Melhado VC, Alvares BR, Almeida OJ. Correlação radiológica e histológica de lesões mamárias não-palpáveis em pacientes submetidas a marcação pré-cirúrgica, utilizando-se o sistema BI-RADS. *Radiol bras*. 2007;40:9-11.
- Sickles EA, D'Orsi CJ, Bassett LW. ACR BI-RADS® Breast Imaging Reporting and Data System. 5 ed. Reston, VA: American College of Radiology; 2013.
- Gebrim LH, Quadros LGA. Rastreamento do câncer de mama no Brasil. *Revista Brasileira de Ginecologia e Obstetrícia*. 2006;28(6):319-23.
- Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Annals of internal medicine*. 2002;137(5 Part 1):347-60.
- Nicandro CR, Efrén MM, María Yaneli AA, Enrique MD, Héctor Gabriel AM, Nancy PC, et al. Evaluation of the diagnostic power of thermography in breast cancer using Bayesian network classifiers. *Computational and mathematical methods in medicine*. 2013;2013:264246.
- West D, West V. Model selection for a medical diagnostic decision support system: a breast cancer detection case. *Artificial intelligence in medicine*. 2000;20(3):183-204.
- Justino JR, Freire FHMdA, Lucio PS. Estimativa de sub-registros de óbitos em pequenas áreas com os métodos bayesiano empírico e algoritmo EM. *Revista Brasileira de Estudos de População*. 2012;29(1):87-100.
- Kahn CE, Jr., Roberts LM, Wang K, Jenks D, Haddawy P. Preliminary investigation of a Bayesian network for mammographic diagnosis of breast cancer. *Proceedings Symposium on Computer Applications in Medical Care*. 1995:208-12.
- Burnside ES, Rubin DL, Shachter RD. Using a Bayesian network to predict the probability and type of breast cancer represented by microcalcifications on mammography. *Studies in health technology and informatics*. 2004;107(Pt 1):13-7.
- Burnside ES, Davis J, Chhatwal J, Alagoz O, Lindstrom MJ, Geller BM, et al. Probabilistic computer model developed from clinical data in national mammography database format to classify mammographic findings. *Radiology*. 2009;251(3):663-72.
- Osuch JR, Reeves MJ, Pathak DR, Kinchelov T. BREASTAID: Clinical results from early development of a clinical decision rule for palpable solid breast masses. *Annals of surgery*. 2003;238(5):728-37.