



PersonalizedUS: Avaliação Personalizada do Risco de Câncer de Mama com Incertezas Interpretáveis

Thiago Rodrigo Ramos (UFSCar)

PersonalizedUS: Interpretable Breast Cancer Risk Assessment with Local Coverage Uncertainty Quantification

Alek Fröhlich^{1*}, Thiago Ramos^{2*}, Gustavo Cabello³, Isabela Buzatto³, Rafael Izbicki², Daniel Tiezzi³

¹UFSC, Florianópolis, Brazil

²UFSCar, São Carlos, Brazil

³USP, Ribeirão Preto, Brazil

alek.frohlich@posgrad.ufsc.br, thiagorr@ufscar.br, gustavocabello@usp.br, ipcarlotti@hcrp.usp.br, rizbicki@ufscar.br, dtiezzi@usp.br

Motivação

O Desafio no Diagnóstico de Câncer de Mama

- Câncer mais comum em mulheres, com 2,3 milhões de novos casos e quase 700 mil mortes em 2022 (Bray et al., 2024).
- Mamografia tem alta taxa de falsos positivos (Ho et al., 2022):
 - 61% das mulheres rastreadas têm ao menos um falso positivo em 10 anos.
 - Até 9% passam por biópsias desnecessárias.
- Consequências:
 - Sobrecarga no sistema de saúde devido ao uso excessivo de recursos médicos.
 - Impacto psicológico significativo para pacientes e famílias.

Ultrassom (US)



- US é amplamente usado como ferramenta auxiliar de diagnóstico (Berg, 2008).
- **Vantagens:** menor custo, avaliação em tempo real e sem radiação (Feig, 2010).
- **Problema:** diagnósticos dependem da experiência do operador, adicionando um grau de incerteza.
- O uso de ultrassom no rastreamento de câncer de mama aumenta em 4-8% o número de pacientes submetidos a biópsias em comparação com a mamografia sozinha, mas apenas 7-8% dessas biópsias guiadas por ultrassom são malignas (Yang et al., 2020; Corsetti et al., 2011).




Sistema BI-RADS

- Sistema **Breast Imaging Reporting and Data System** é obtido a partir da realização de exames de imagem como mamografia, ultrassonografia e ressonância magnética
- Decisões baseadas no BI-RADS:
 - Radiologistas classificam lesões como benignas, necessitando acompanhamento ou biópsia com base no sistema BI-RADS (Mendelson et al., 2013).
- Limitações do BI-RADS:
 - Pouca orientação para subdivisões das categorias 4a, 4b e 4c, que apresentam risco variando de 2%-95% (Pfob et al., 2022).
 - Alta variabilidade, contribuindo para taxas elevadas de falsos positivos (Lazarus et al., 2006).

Final Assessment Categories

Category	Management	Likelihood of cancer	
0	Need additional imaging or prior examinations	Recall for additional imaging and/or await prior examinations	n/a
1	Negative	Routine screening	Essentially 0%
2	Benign	Routine screening	Essentially 0%
3	Probably Benign	Short interval-follow-up (6 month) or continued	>0% but ≤ 2%
4	Suspicious	Tissue diagnosis	4a. low suspicion for malignancy (>2% to ≤ 10%) 4b. moderate suspicion for malignancy (>10% to ≤ 50%) 4c. high suspicion for malignancy (>50% to <95%)
5	Highly suggestive of malignancy	Tissue diagnosis	≥95%
6	Known biopsy-proven	Surgical excision when clinical appropriate	n/a


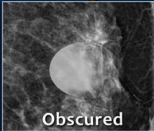
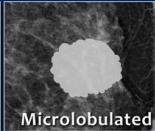
Shape

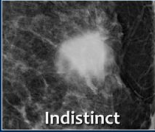
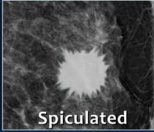




The shape of a mass is either round, oval or irregular.

Always make sure that a mass that is found on physical examination is the same as the mass that is found with mammography or ultrasound.
Location and size should be applied in any lesion, that must undergo biopsy.

Margin

The margin of a lesion can be:

- Circumscribed (historically well-defined). This is a benign finding.
- Obscured or partially obscured, when the margin is hidden by superimposed fibroglandular tissue. Ultrasound can be helpful to define the margin better.
- Microlobulated. This implies a suspicious finding.
- Indistinct (historically ill-defined). This is also a suspicious finding.
- Spiculated with radiating lines from the mass is a very suspicious finding.

Objetivos



- **Objetivos**
 - Reduzir falsos positivos no padrão BI-RADS sem comprometer casos de câncer.
 - Desenvolver um modelo interpretável para suporte médico.
 - Fornecer **quantificação personalizada de incertezas** com garantias teóricas.
- **Nossa solução**
 - Sistema baseado em regressão logística integrado a conformal prediction.
 - Adaptação do método **Locart** (Cabezas et al., 2024) para fornecer quantificação de incerteza personalizada.



Dados

Nossos Dados



- 1936 lesões mamárias com dados clínicos, BI-RADS e características Doppler.
- Estudo multicêntrico observacional com dados retrospectivos e prospectivos.
 - Retrospec.: HCRP-USP (2014–2021) e MATER (2019–2021).
 - Prospec.: HCRP-USP, MATER, Hospital de Amor (Barretos e Campo Grande) desde 2021.
- Divisão dos Dados:
 - Treinamento: 513 (apenas dados retrospectivos).
 - **Calibração: 1059 (dados mistos).**
 - Teste: 364 (apenas dados prospectivos).

	age	size	palpable	birads	vessels	ri	shape	margins	orientation	affiliation	us_date	study
712	51.0	17.0	1.0	4a	0.0	0.00	oval	microlobulated	parallel	hcrp	NaN	retrospective
210	39.0	20.0	1.0	4b	0.0	0.00	irregular	indistinct	not parallel	mater	10/01/20	retrospective
702	43.0	9.0	0.0	4a	0.0	0.00	oval	circumscribed	parallel	hcrp	NaN	retrospective
506	78.0	12.0	0.0	4b	1.0	1.00	irregular	microlobulated	parallel	hcrp	NaN	retrospective
525	41.0	17.0	0.0	4a	1.0	1.00	irregular	spiculated	not parallel	hcrp	NaN	retrospective
...
717	65.0	8.2	0.0	4c	1.0	0.71	irregular	indistinct	not parallel	hcrp	NaN	retrospective
773	50.0	26.0	1.0	4c	1.0	0.77	irregular	indistinct	parallel	hcrp	NaN	retrospective
847	50.0	19.0	1.0	5	1.0	0.69	irregular	indistinct	not parallel	hcrp	NaN	retrospective
568	40.0	22.0	1.0	4a	1.0	1.00	oval	circumscribed	parallel	hcrp	NaN	retrospective
693	60.0	29.0	1.0	4c	0.0	0.00	irregular	indistinct	not parallel	hcrp	NaN	retrospective

Prospective Stats	Age	Size	Palpable	Vessels	RI
Count	909.000000	909.000000	909.000000	909.000000	909.000000
Mean	51.952695	16.697360	0.551155	0.524752	0.395831
Std	13.577278	6.994456	0.497650	0.499662	0.388893
Min	19.000000	1.000000	0.000000	0.000000	0.000000
25%	42.000000	11.000000	0.000000	0.000000	0.000000
50%	52.000000	16.000000	1.000000	1.000000	0.510000
75%	62.000000	23.000000	1.000000	1.000000	0.760000
Max	93.000000	30.000000	1.000000	1.000000	1.000000

Retrospective Stats	Age	Size	Palpable	Vessels	RI
Count	1027.000000	1027.000000	1027.000000	1027.000000	1027.000000
Mean	51.436222	16.735638	0.641675	0.536514	0.423671
Std	15.178360	6.854130	0.479742	0.498908	0.410031
Min	18.000000	3.000000	0.000000	0.000000	0.000000
25%	41.000000	11.000000	0.000000	0.000000	0.000000
50%	51.000000	16.000000	1.000000	1.000000	0.570000
75%	61.000000	22.000000	1.000000	1.000000	0.800000
Max	92.000000	30.000000	1.000000	1.000000	1.300000

Prospective Proportions - Shape Proportion

Irregular	0.611661
Oval	0.341034
Round	0.047305

Prospective - Margins Proportion

Angular	0.201320
Circumscribed	0.268427
Indistinct	0.135314
Microlobulated	0.157316
Spiculated	0.237624

Prospective - Orientation Proportion

Not Parallel	0.261826
Parallel	0.738174

Retrospective - Shape Proportion

Irregular	0.518014
Oval	0.430380
Round	0.051607

Retrospective - Margins Proportion

Angular	0.058423
Circumscribed	0.365141
Indistinct	0.262902
Microlobulated	0.158715
Spiculated	0.154820

Retrospective - Orientation Proportion

Not Parallel	0.32814
Parallel	0.67186



PersonalizedUS

PersonalizedUS

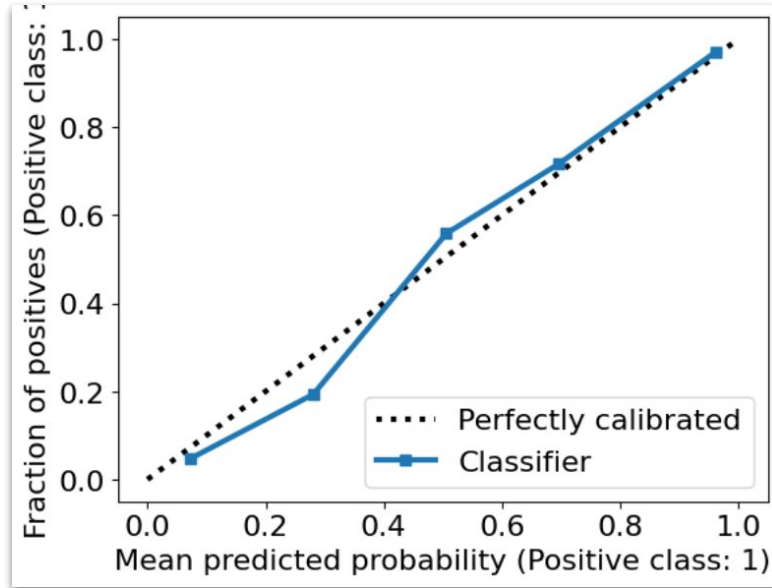


- Sistema baseado em regressão logística integrado a conformal prediction.
- Adaptação do método **Locart** (Cabezas et al., 2024) para fornecer quantificação de incerteza personalizada.

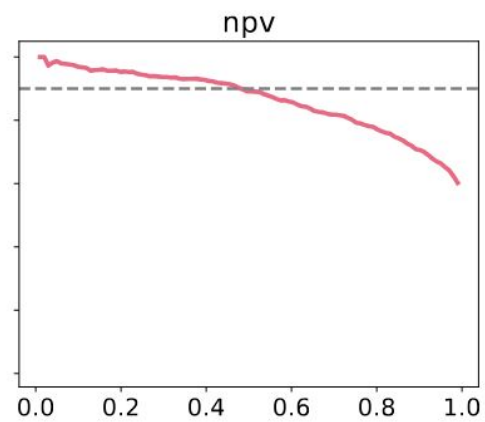
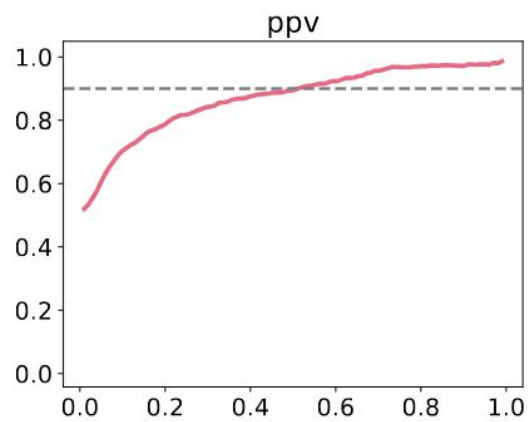
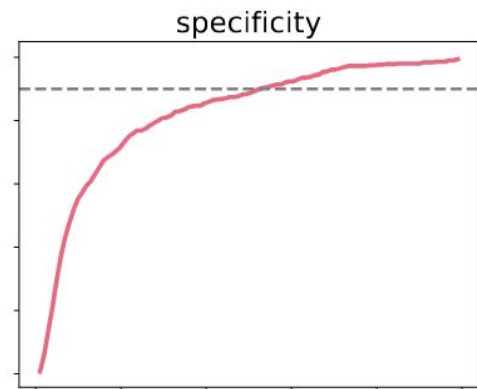
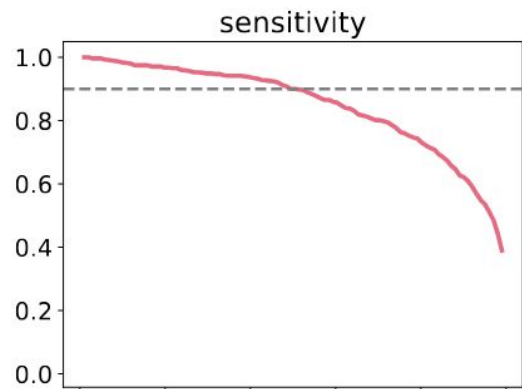
Modelo de Avaliação de Risco



- Regressão Logística:
 - Escolhido pela simplicidade e interpretabilidade.
 - Permite que médicos compreendam como cada variável contribui para a previsão.
 - Não há vantagem significativa no uso de modelos mais complexos, considerando o custo da interpretabilidade.
 - A regressão logística oferece um excelente equilíbrio entre desempenho, simplicidade e aplicabilidade clínica.



- Divisão dos Dados:
 - Treinamento: 513 (apenas dados retrospectivos).
 - **Calibração: 1059 (dados mistos).**
 - Teste: 364 (apenas dados prospectivos).



Model	Sensitivity	Specificity	PPV	NPV
Logistic	0.900	0.888	0.897	0.892
DT	0.770	0.947	0.940	0.793
NB	0.687	0.975	0.967	0.743
KNN	0.903	0.894	0.902	0.896
XGB	0.865	0.871	0.878	0.857
NN	0.885	0.904	0.908	0.880
AB	0.818	0.908	0.905	0.822
RF	0.900	0.878	0.888	0.891
SVM	0.898	0.896	0.903	0.891

- **Sensibilidade:** Proporção de verdadeiros positivos entre os que realmente têm a condição.
- **PPV (Valor Preditivo Positivo):** Proporção de verdadeiros positivos entre os resultados positivos do teste.

Quantificação de Incerteza com Locart

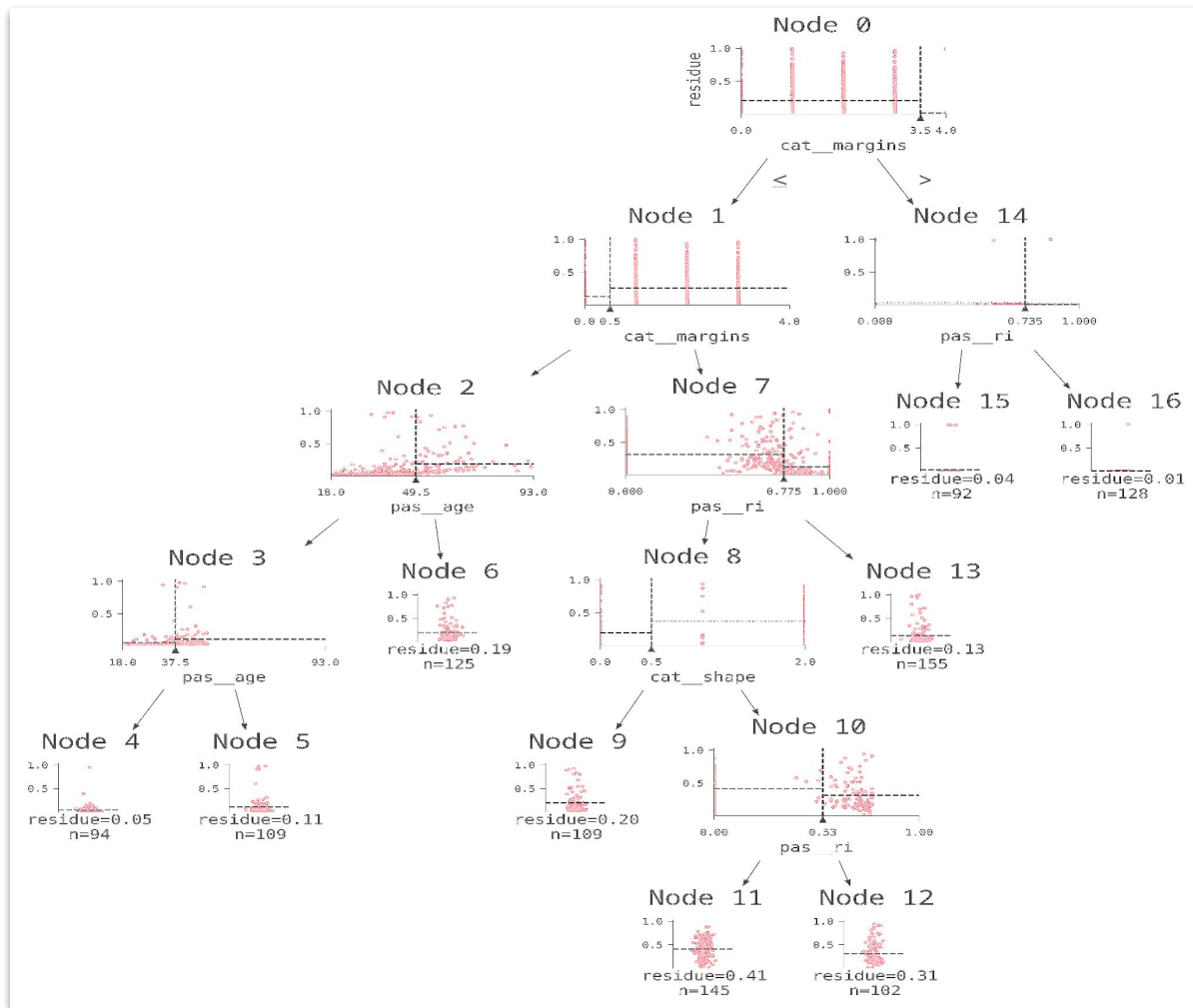


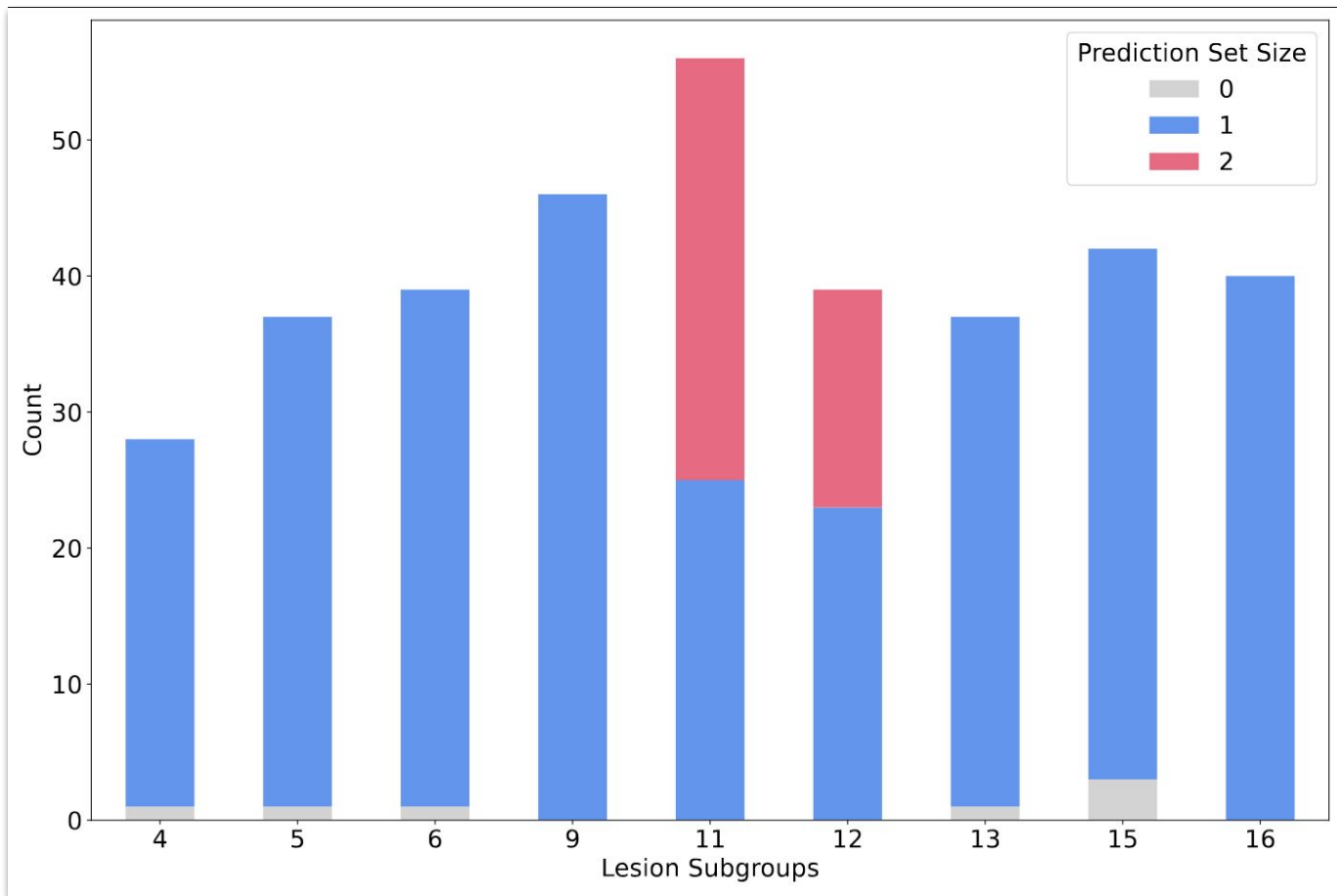
- Objetivo
 - Fornecer conjuntos de previsão ajustados à complexidade de cada lesão.
- Como Funciona
 - Cálculo dos Resíduos: Discrepâncias entre previsões do modelo e resultados observados.
 - Particionamento: Árvore de decisão divide o espaço das lesões em subgrupos com características semelhantes.
 - Conjuntos de Previsão: Conjuntos de Incerteza personalizados são calculados para cada subgrupo, considerando a dificuldade local.
 - Esses conjuntos são da forma: vazio, {Câncer}, {Não-Cancer}, {Cancer, Não-Câncer}
 - **Intuitivamente:** Com chance de 90% o valor real está no respectivo conjunto.

Quantificação de Incerteza com Locart



- Objetivo
 - Fornecer conjuntos de previsão ajustados à complexidade de cada lesão.
- Como Funciona
 - Aprendemos a tarefa de previsão de Câncer via Regressão Logística no conjunto de **treino**.
 - No conjunto de **calibração**, utilizamos uma árvore de decisão para dividir o espaço das lesões em subgrupos com características semelhantes.
 - Também no conjunto de **calibração**, os conjuntos de Incerteza personalizados são calculados para cada subgrupo, considerando a dificuldade local.
 - Esses conjuntos são da forma: vazio, {Câncer}, {Não-Cancer}, {Cancer, Não-Câncer}
 - **Intuitivamente:** Com chance de 90% o valor real está no respectivo conjunto.





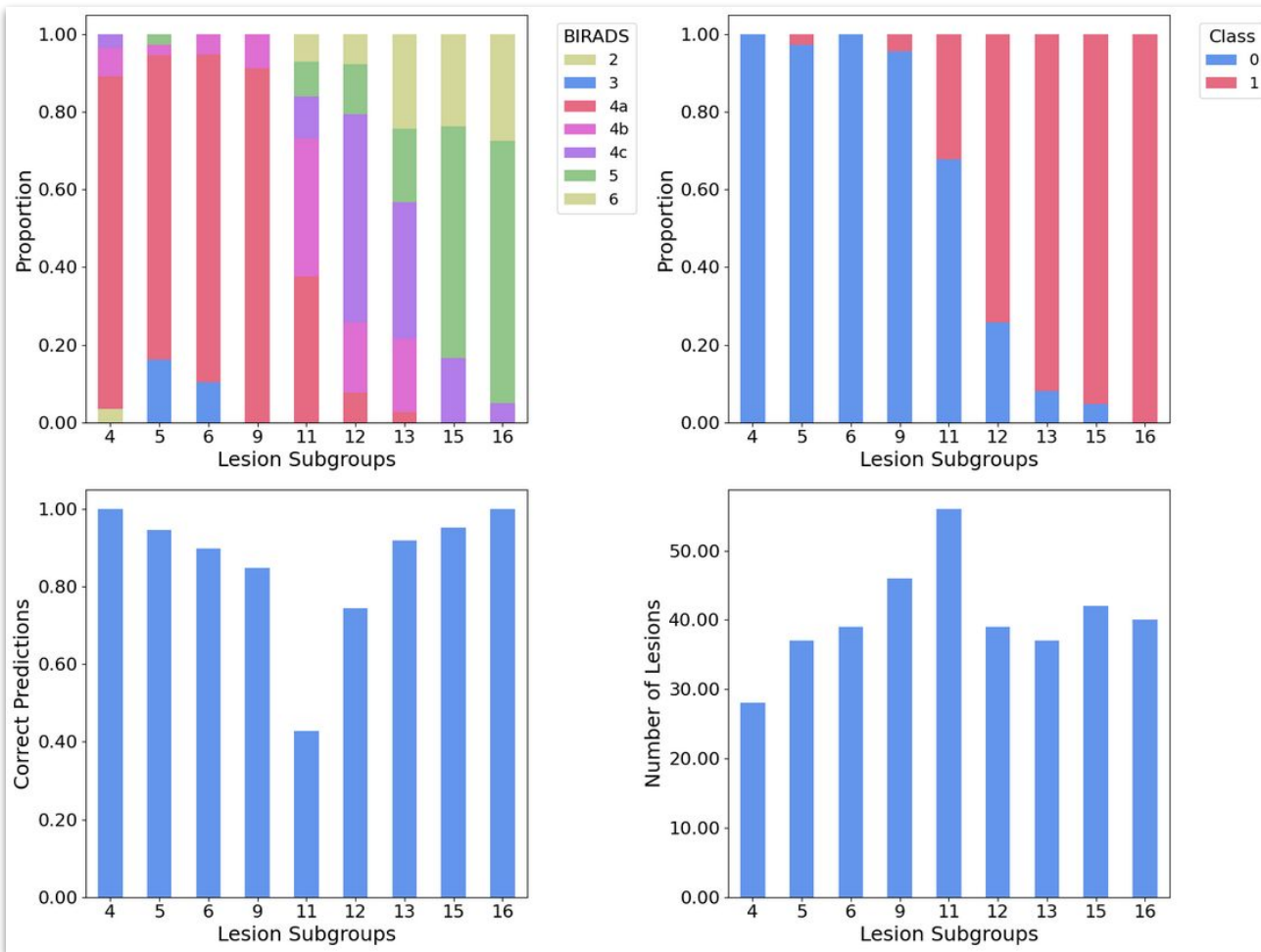
Subgrupos de Lesões

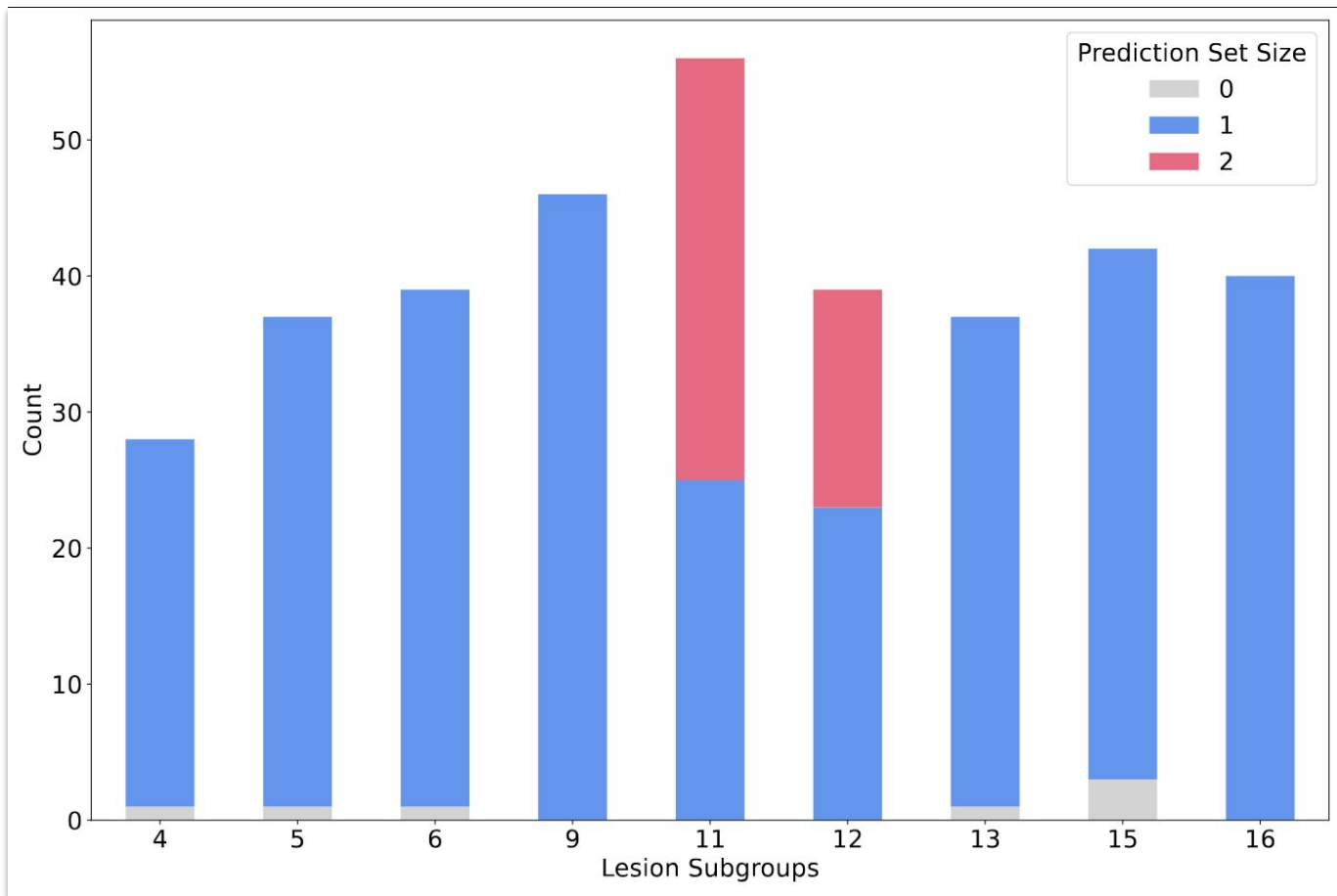
Avaliação da Quantificação de Incertezas



- **Critérios Avaliados:**
 - Tamanho médio dos conjuntos.
 - cobertura empírica.
 - proporção de conjuntos contendo apenas o valor verdadeiro.
- A maioria dos subgrupos apresentou conjuntos de previsão próximos de um elemento, exceto os mais desafiadores (subgrupos 11 e 12).
- Lesões desafiadoras representam 20% do conjunto de teste, mas os outros 80% tiveram previsões confiáveis.

Leaf	Avg. Set Size	Emp. Cov. (%)	Truth Only (%)
4	0.96	96.43	96.43
5	0.97	94.59	94.59
6	0.97	97.44	97.44
9	1.00	95.65	95.65
11	1.55	89.29	33.93
12	1.41	94.87	53.85
13	0.97	89.19	89.19
15	0.93	90.48	90.48
16	1.00	100.00	100.00





Conclusão



- Redução de Biópsias Desnecessárias:
 - Em lesões BI-RADS 4a e 4b, o modelo recomendaria apenas 64 de 196 biópsias, uma redução de mais de 65%.
- Eficácia Diagnóstica:
 - Nenhum caso de câncer seria perdido com o limiar otimizado ($t=0.2$).
- Impacto Prático:
 - Otimização do uso de recursos médicos.
 - Menor estresse para pacientes e famílias, com foco em diagnósticos mais precisos e ações prioritárias.

Muito obrigado!

