



Centro para Inteligência Artificial em Gestão de Saúde  
<https://ciagsaude.icmc.usp.br/pt-BR>



# Do Diagnóstico à Prevenção: Como a IA Está Antecipando Riscos na Saúde

Paulo Mazzoncini de Azevedo-Marques E.E Ph.D.

Departamento de Imagens Médicas, Hematologia e Oncologia Clínica

Faculdade de Medicina de Ribeirão Preto - USP

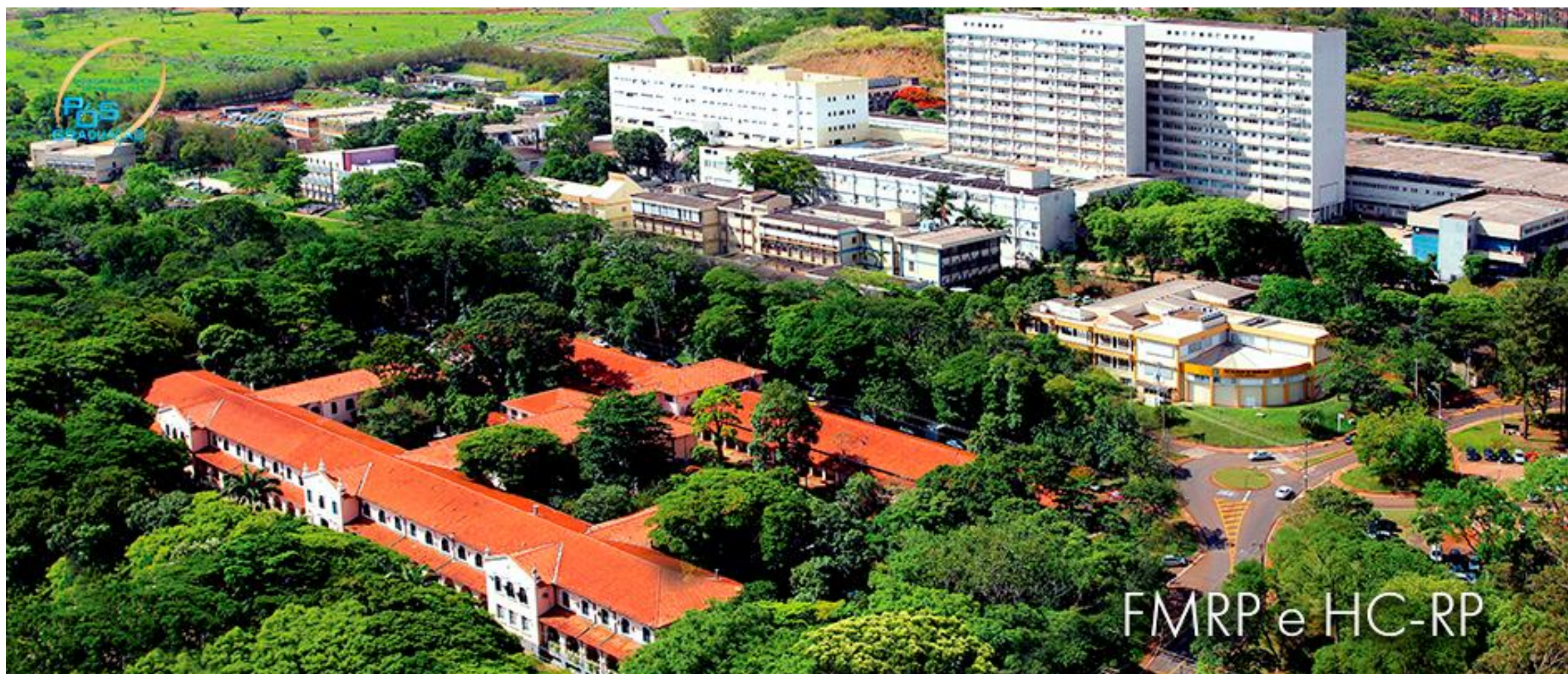
[pmarques@fmrp.usp.br](mailto:pmarques@fmrp.usp.br)











FMRP e HC-RP





**Centro para Inteligência Artificial em Gestão de Saúde**  
<https://ciagsaude.icmc.usp.br/pt-BR>

## Research & Innovation Leaders



## Directorate of Strategy



**160**  
Health & AI Investigators

# Main concepts



## Intelligent Healthcare Pathways

Translational Research by IA

Cost-effective solutions

a proper tradeoff



# Conteúdo Abordado

- Diagnóstico Auxiliado por Computador (CAD-AI)
  - CADt radiografia de tórax
- Radiômica
  - Caracterização de câncer primário de pulmão
    - Cavidade vs Pseudocavidade
    - Sobrevida vs heterogeneidade tumoral
    - Predição de metástase distante
- Previsão de eventos cardiovasculares usando aprendizado de máquina

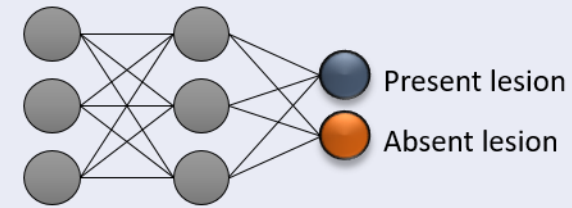


# APLICAÇÕES CLÍNICAS DE INTELIGÊNCIA ARTIFICIAL PREDITIVA EM IMAGENS MÉDICAS

**CAD**

década de 80

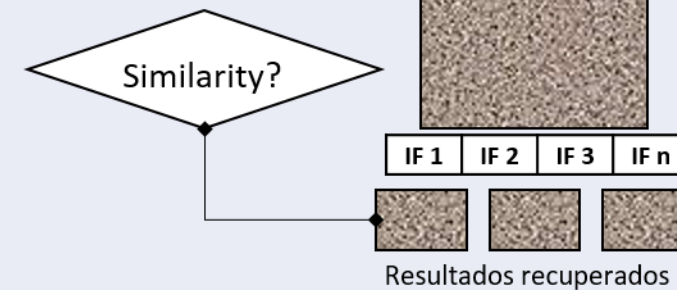
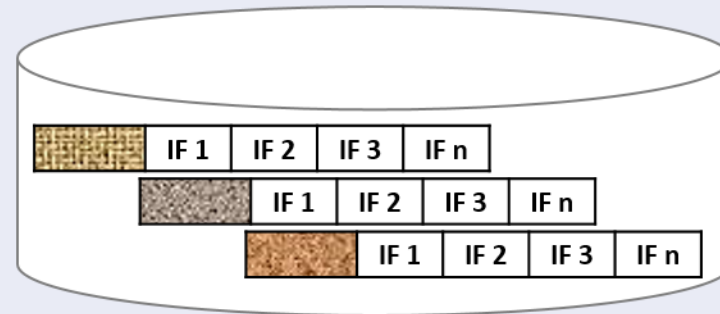
Patient ID	Image Features			
	IF 1	IF 2	IF 3	IF n
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...



**CBIR**

anos 2000

(hospital digital)

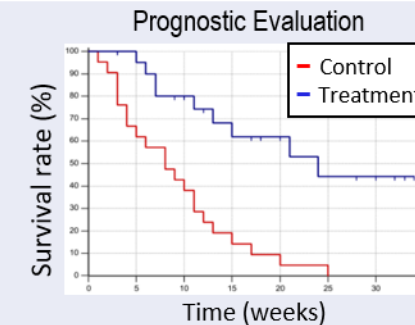


**Radiomics**

2012- ...

(big data)

Patient ID	Image Features			Clinical Data		Genomic Data	
	IF 1	IF 2	IF n	CD 1	CD n	GD 1	GD n
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...



# DIAGNÓSTICO AUXILIADO POR COMPUTADOR

## *Computer-aided diagnosis (CAD)*

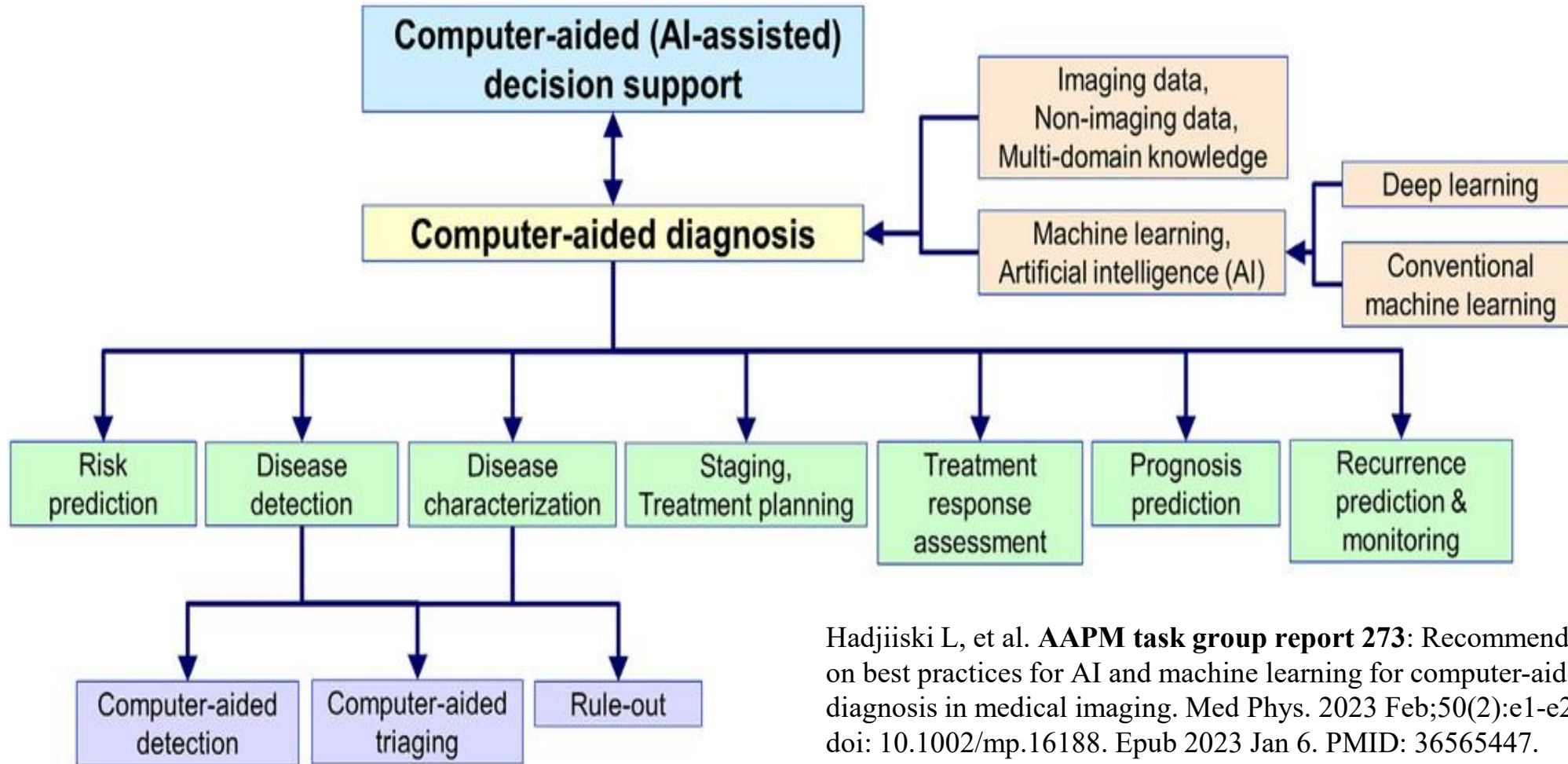
- Diagnóstico feito por especialista que utiliza a saída de uma análise computadorizada de imagens como **“segunda opinião”** no processo de detecção de lesões e tomadas de decisões diagnósticas.

Doi K. Computer-aided diagnosis and its potential impact on diagnostic radiology. In: Doi k, MacMahon H, Giger ML, Hoffmann KR, eds. Computer-aided diagnosis in medical imaging. Amsterdam: Elsevier Science, 1999:11-20.



# CAD-AI

**REFERE-SE A UM AMBIENTE EXPANDIDO DE SUPORTE À DECISÃO CLÍNICA QUE USA MÉTODOS DE IA TRADICIONAIS E BASEADOS EM DL.**



Hadjiiski L, et al. **AAPM task group report 273**: Recommendations on best practices for AI and machine learning for computer-aided diagnosis in medical imaging. Med Phys. 2023 Feb;50(2):e1-e24. doi: 10.1002/mp.16188. Epub 2023 Jan 6. PMID: 36565447.

## Redes Neurais Convolucionais na Classificação de Imagens de RX de Tórax: COVID-19, Doenças Intersticiais Pulmonares e Tuberculose

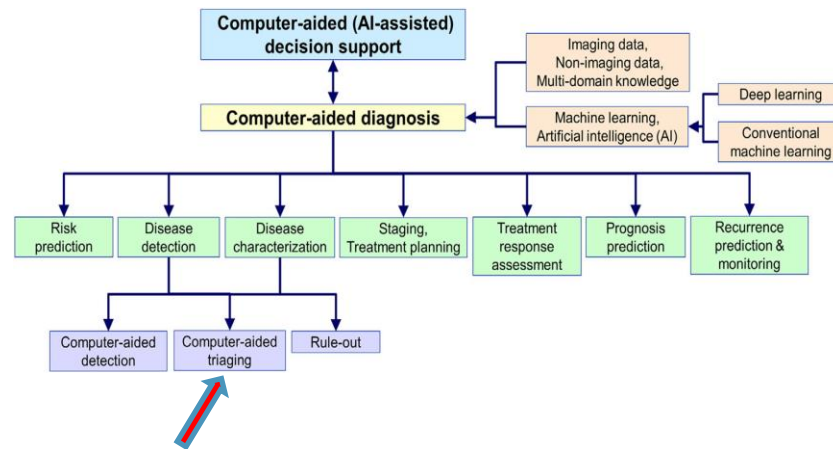


Tabela 3 – Quantidade de dados utilizados em cada banco

Banco	HCFMRP	Curated COVID-19	MC	Shenzhen	total
covid19	189	370	0	0	559
dip	308	0	0	0	308
normais	382	177	0	0	559
tuberculose	165	0	58	336	559

Fonte: Autor.

LIMA, L. L. d. et al. HCFMRP COVID-19 & LID (v1). Zenodo, 2022. Disponível em: <https://doi.org/10.5281/zenodo.6841440>.

FRANCIS, M. Curated COVID-19 Chest X-Ray Dataset. 2020. <https://www.kaggle.com/datasets/francismon/curated-covid19-chest-xray-dataset>.

JAEGER, S. et al. Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. Quantitative Imaging in Medicine and Surgery, AME Publishing Company, v. 4, n. 6, p. 475–477, 2014.

### Apoio Financeiro:

FAPESP #16/17078-0 #14/50889-7

CNPq #305124/2018-8 #302480/2022-6

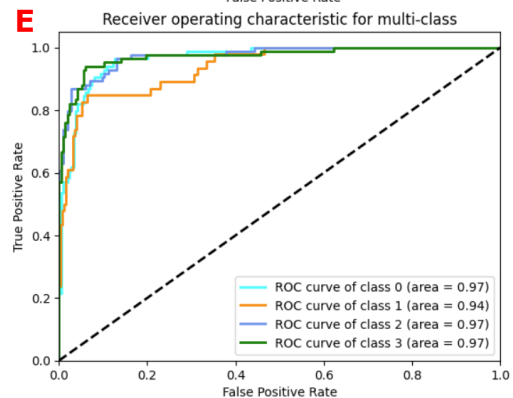
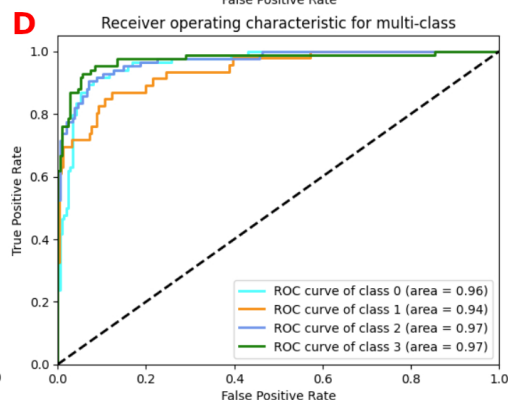
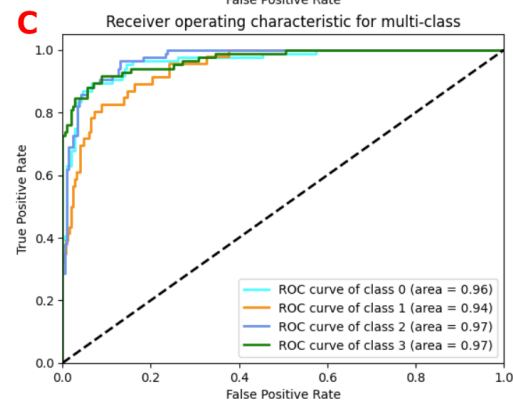
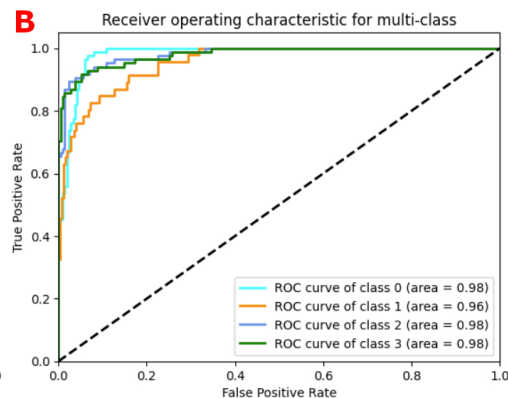
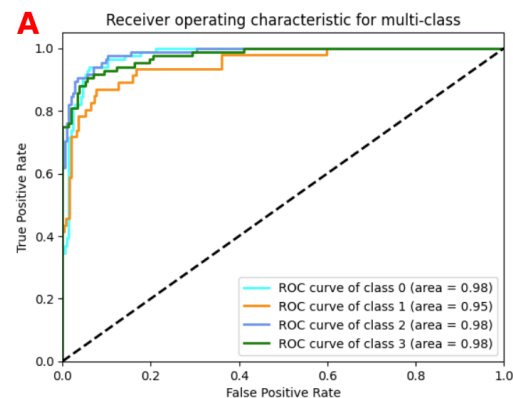
CAPES #001

Pesquisa aprovada pelo CEP-HCRP  
com dispensa de TCLE

Mauricio Koenigkam Santos. Redes Neurais Convolucionais na Classificação de Imagens de RX de Tórax: COVID-19, Doenças Intersticiais Pulmonares e Tuberculose. Dissertação. Bioengenharia. USP. 2023.

<https://github.com/mauks37/CNN-Chest-X-Ray>





Curvas ROC e valores de AUC-ROC do experimento com as RNCs, grupo teste: A- VGG19, B- EfficientNetV2S, C- ResNet152V2, D- InceptionV3 e E- Xception.

Class 0 – covid  
 Class 1 - dip  
 Class 2 - normal  
 Class 3 – tuberculose

Matriz de Confusão e Relatório do experimento com as RNCs, grupo teste: VGG19, EfficientNetV2S, ResNet152V2, InceptionV3 e Xception.

		PREVISÃO VGG19							
		covid	dip	normal	tuberculose	covid	precisão	sensibilidade	f1-score
REAL	covid	78	2	3	1	covid	0,87	0,93	0,90
	dip	6	35	1	4	dip	0,80	0,76	0,78
	normal	3	2	76	3	normal	0,93	0,90	0,92
	tuberculose	3	5	2	74	tuberculose	0,90	0,88	0,89
acurácia							88,26%		

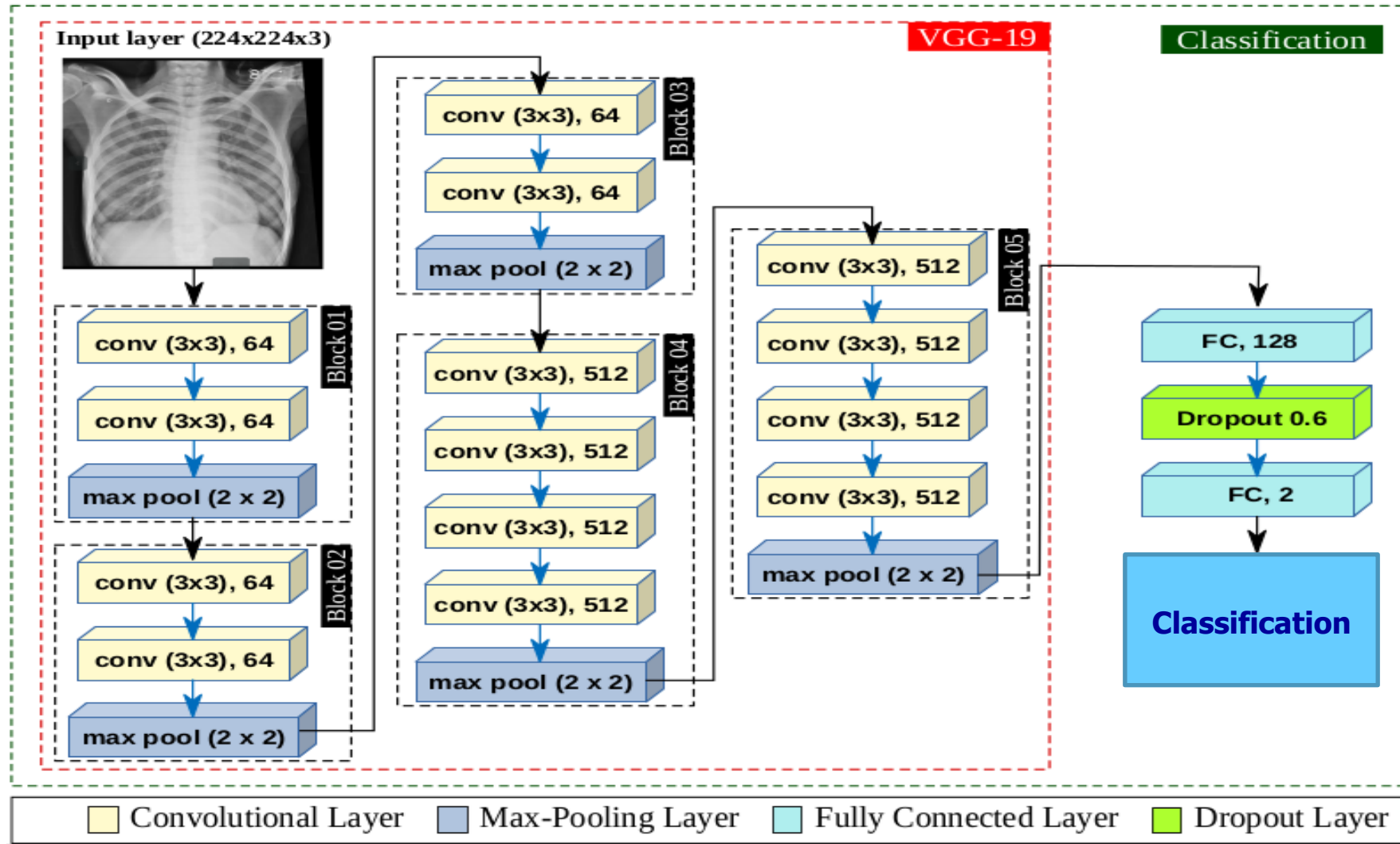
		PREVISÃO EfficientNetV2S							
		covid	dip	normal	tuberculose	covid	precisão	sensibilidade	f1-score
REAL	covid	75	3	5	1	covid	0,86	0,89	0,88
	dip	5	35	2	4	dip	0,73	0,76	0,74
	normal	4	3	74	3	normal	0,91	0,88	0,90
	tuberculose	3	7	0	74	tuberculose	0,90	0,88	0,89
acurácia							86,58%		

		PREVISÃO ResNet152V2							
		covid	dip	normal	tuberculose	covid	precisão	sensibilidade	f1-score
REAL	covid	75	2	5	2	covid	0,84	0,89	0,87
	dip	8	32	1	5	dip	0,73	0,70	0,71
	normal	2	7	70	5	normal	0,89	0,83	0,86
	tuberculose	4	3	3	74	tuberculose	0,86	0,88	0,87
acurácia							84,23%		

		PREVISÃO InceptionV3							
		covid	dip	normal	tuberculose	covid	precisão	sensibilidade	f1-score
REAL	covid	74	3	7	0	covid	0,84	0,88	0,86
	dip	7	33	3	3	dip	0,79	0,72	0,75
	normal	4	1	72	7	normal	0,86	0,86	0,86
	tuberculose	3	5	2	74	tuberculose	0,88	0,88	0,88
acurácia							84,90%		

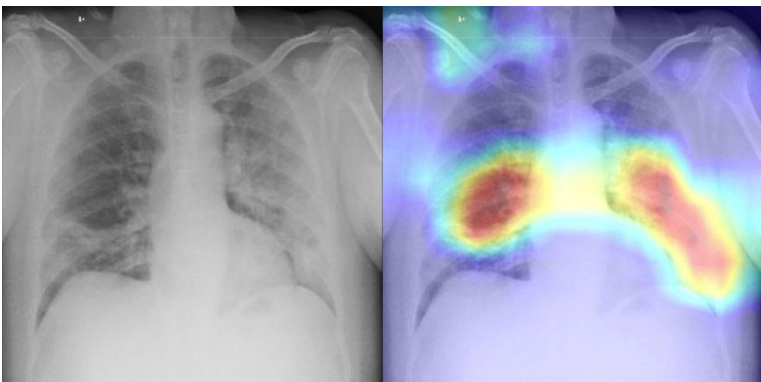
		PREVISÃO Xception							
		covid	dip	normal	tuberculose	covid	precisão	sensibilidade	f1-score
REAL	covid	71	3	9	1	covid	0,83	0,85	0,84
	dip	6	30	6	4	dip	0,77	0,65	0,71
	normal	5	0	75	4	normal	0,80	0,89	0,84
	tuberculose	4	6	4	70	tuberculose	0,89	0,83	0,86
acurácia							82,55%		

# Arquitetura Modificada (Transfer Learning) VGG - 19

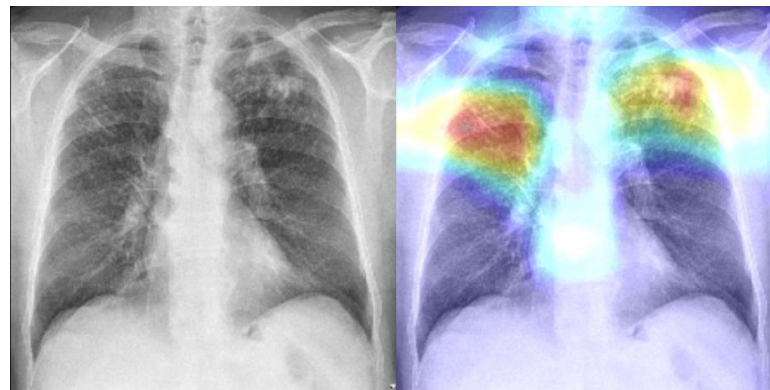




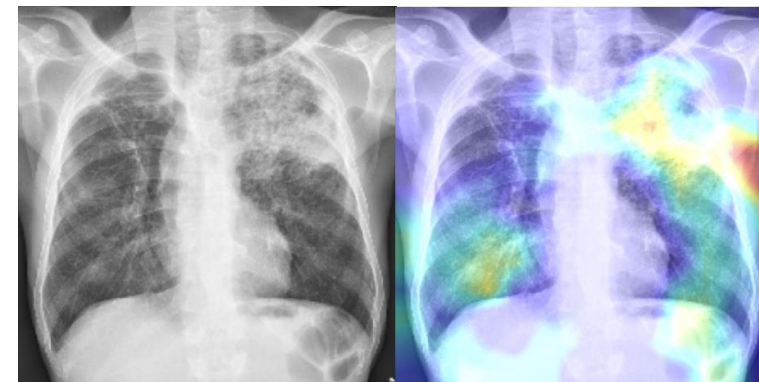
# Resultados (interpretabilidade)



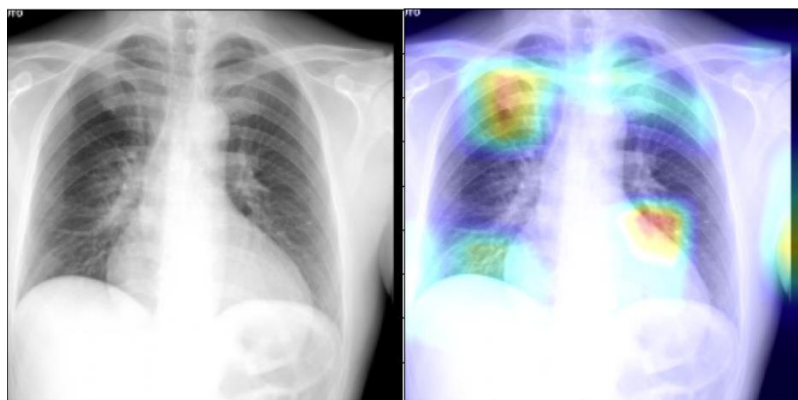
Comparação de um RXT do grupo teste, com COVID-19 (esquerda) com o **Grad-CAM Heatmap** (direita). O modelo deu importância à área de interesse correta e acertou a classificação.



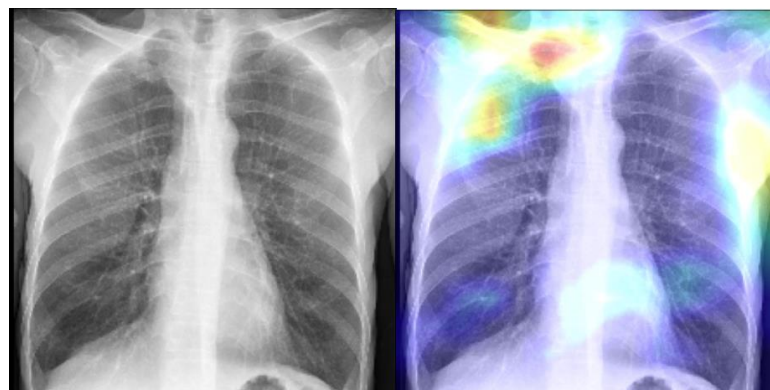
Comparação de um RXT do grupo teste, com DIP granulomatosa (esquerda) com o **Grad-CAM Heatmap** (direita). O modelo deu importância à área de interesse correta e acertou a classificação.



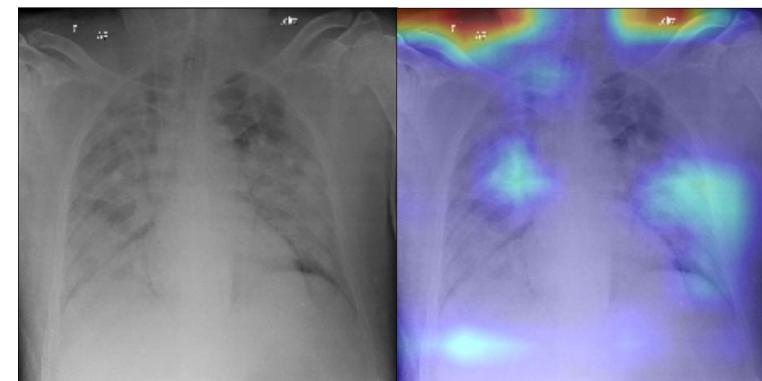
Comparação de um RXT do grupo teste, com tuberculose (esquerda) com o **Grad-CAM Heatmap** (direita). O modelo deu importância à área de interesse correta e acertou a classificação.



RXT do grupo teste, com COVID-19 (esquerda) com o **Grad-CAM Heatmap** (direita). O modelo classificou como 'tuberculose'.



RXT do grupo teste, sem doença pulmonar ('normal') (esquerda) com o **Grad-CAM Heatmap** (direita), que focou principalmente no ápice do pulmão direito. O modelo classificou como 'tuberculose'.

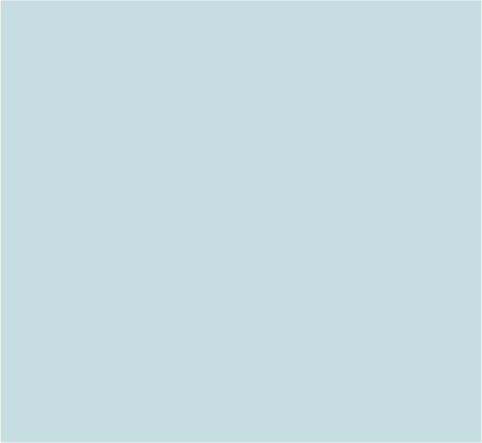


RXT do grupo teste, com COVID-19 (esquerda) com o **Grad-CAM Heatmap** (direita), que focou em artefatos fora do tórax. O modelo classificou corretamente a imagem.

# RADIÔMICA

- ✓ **Extração massiva** de características quantitativas (atributos) de imagens médicas e subsequente reconhecimento de padrões para **detecção/diagnóstico** auxiliado por computador e obtenção de outras informações, como **marcadores prognósticos** e correlação com dados genéticos (genômica).

Radiology



Robert J. Gillies, PhD  
Paul E. Kinahan, PhD  
Hedvig Hricak, MD, PhD, Dr(hc)

**Radiomics:** Images Are More than Pictures, They Are Data<sup>1</sup>

In the past decade, the field of medical image analysis has grown exponentially, with an increased number of pattern recognition tools and an increase in data set sizes. These advances have facilitated the development of processes for

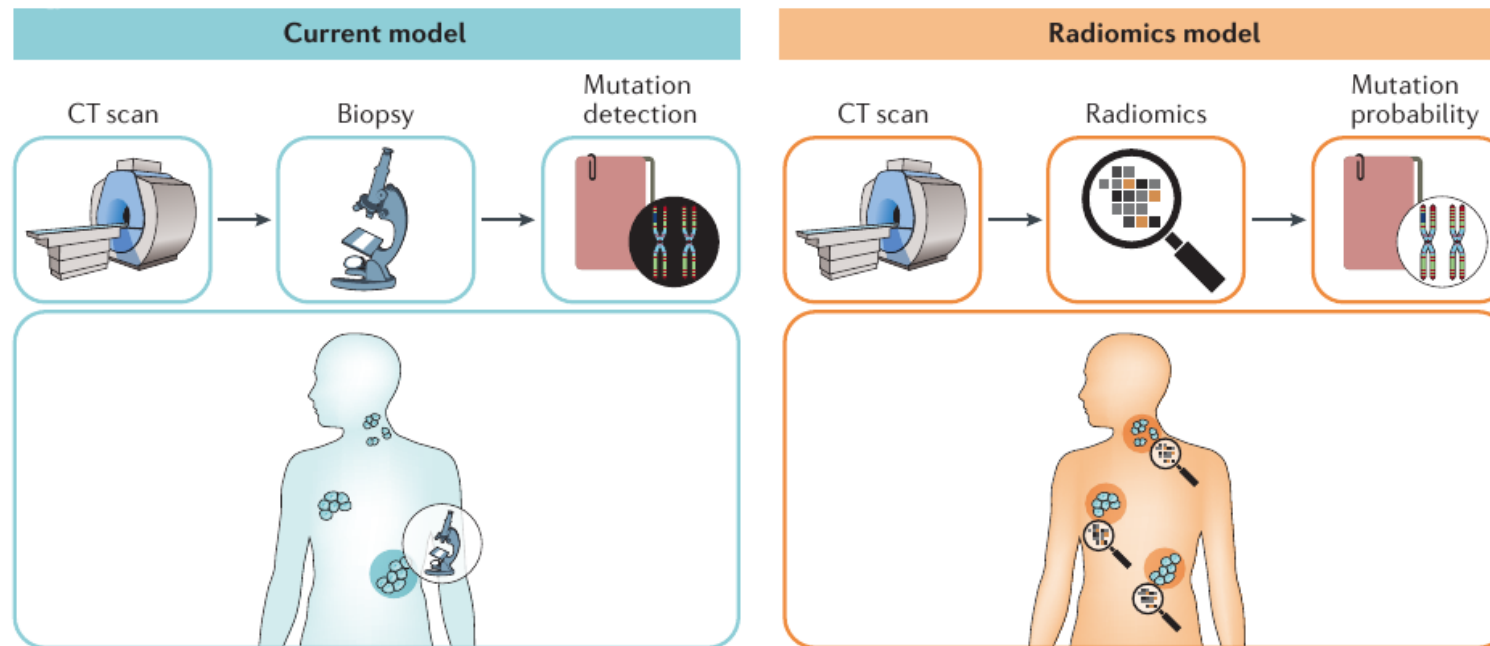
ORIGINAL RESEARCH ■ SPECIAL REPORT

Gillies RJ et al. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016  
doi: 10.1148/radiol.2015151169.

# Radiômica - Modelo

- Radiômica tem se firmado como uma ferramenta eficiente para fornecer suporte diagnóstico e prognóstico para medicina de precisão, a partir de biomarcadores quantitativos extraídos de exames por imagem.***

Ferreira Junior JR, Koenigkam-Santos M, Cipriano FEG, Fabro AT, Azevedo-Marques PM (2018) Radiomics-based features for pattern recognition of lung cancer histopathology and metastases. Computer Methods and Programs in Biomedicine 159:23-30.

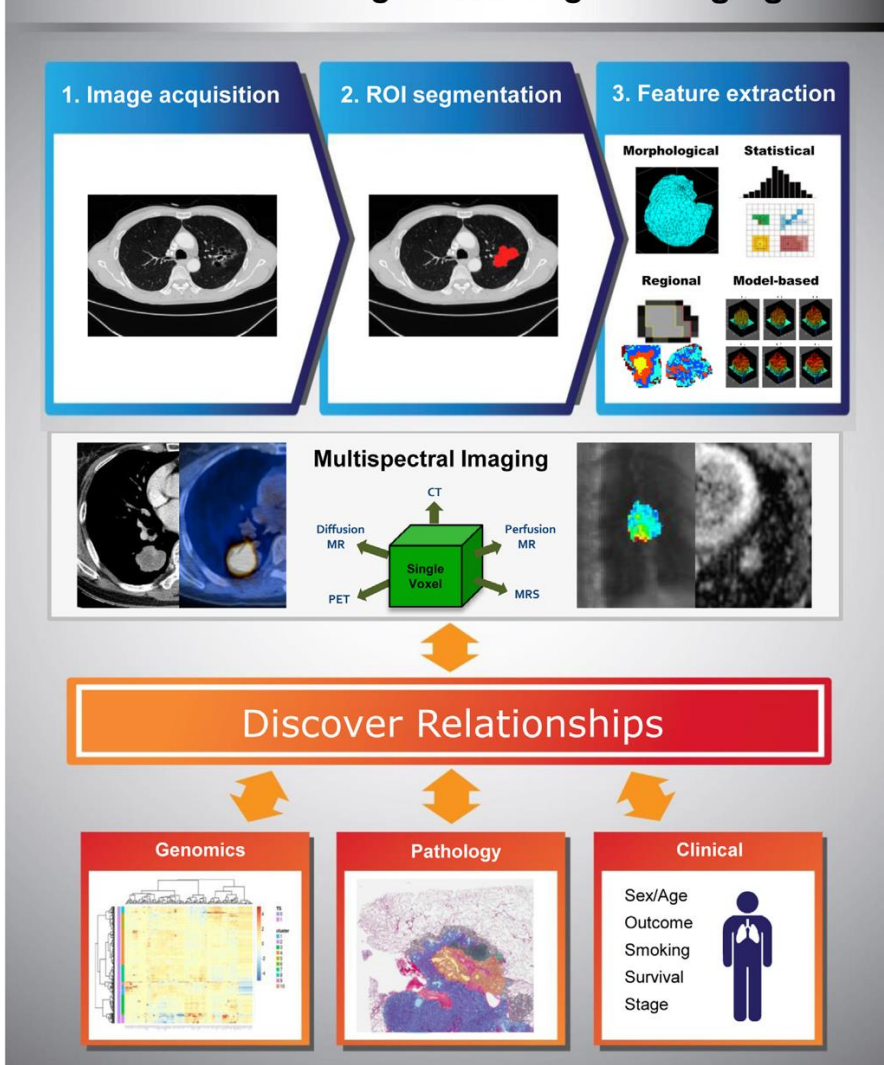


Lambin P, et al. (2017) Radiomics: the bridge between medical imaging and personalized medicine. Nature Reviews Clinical Oncology 14(12):749.



# Radiômica - Etapas

## Radiomics : Processing of Radiological Imaging Data



- ✓ A análise radiômica de atributos de imagem, com base na intensidade dos níveis de cinza, forma, textura, tamanho e volume, pode fornecer informações sobre genótipo tumoral, fenótipo e heterogeneidade clonal.
- ✓ A RADIÔMICA atende ao conceito de medicina de precisão ou personalizada.

Lee G. et al. Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art. EJR. 2017.  
<https://doi.org/10.1016/j.ejrad.2016.09.005>

# EXEMPLOS DE APLICAÇÕES PARA CARACTERIZAÇÃO DO CÂNCER PRIMÁRIO DE PULMÃO

- ✓ Pesquisas aprovadas pelo Comitê de Ética em Pesquisa do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto (CEP-HCRP).
- ✓ Suporte financeiro:
  - ✓ Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) #16/17078-0, #14/50889-7
  - ✓ Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) #305124/2018-8, #302480/2022-6
  - ✓ Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) #001
- ✓ Exames de imagem = ferramentas importantes de diagnóstico, especialmente em câncer.
- ✓ Imagem de pulmão = método mais usado Tomografia Computadorizada (TC).
- ✓ Rotina diária: descrição da localização das lesões pulmonares, tamanho, morfologia, atenuação, realce pós-contraste.
- ✓ “Descrição tradicional” = benigno x maligno .... primário x secundário.
- ✓ Hoje = qual subtipo de câncer é? Estadiamento? Sinais de pior prognóstico? Mutações?

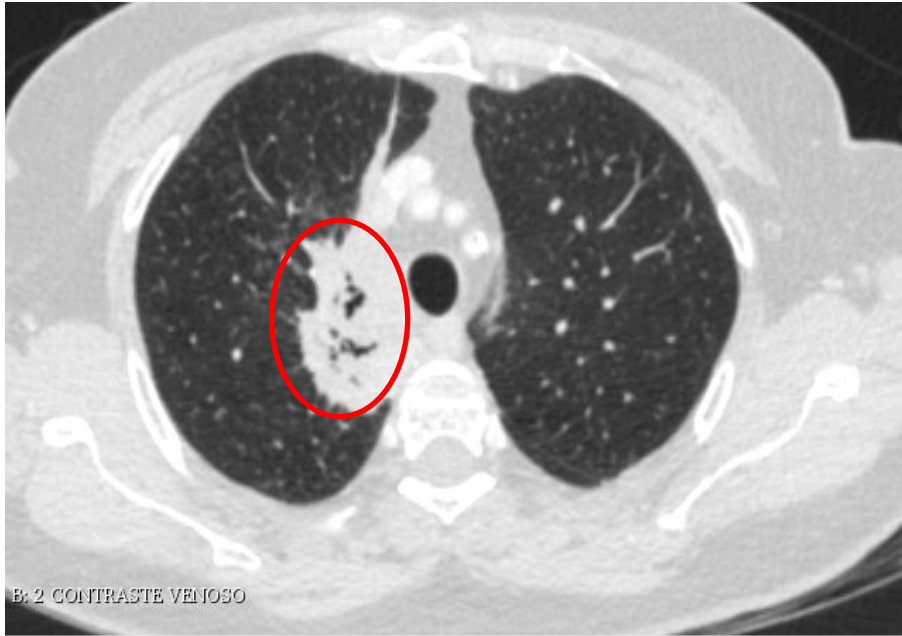
# PRINCIPAIS SUBTIPOS DE CÂNCER PRIMÁRIO DE PULMÃO

- Câncer de pulmão de não pequenas células (CPNPC) – 80% a 85%
  - Adenocarcinoma** – é o tipo mais predominante atualmente. Começa nas células secretoras de substâncias, como o muco. Este tipo de câncer de pulmão ocorre principalmente em fumantes ou ex-fumantes, mas também é o tipo mais comum de câncer pulmonar observado em não-tabagistas. É mais comum em mulheres do que em homens, e tem maior probabilidade de ocorrência em pessoas jovens do que os outros tipos de câncer de pulmão;
  - Carcinoma de células escamosas (ou espinocelular)** – começa nas células escamosas, que são células planas que revestem o interior das vias aéreas. Geralmente está relacionado ao tabagismo e costuma se localizar na parte central dos pulmões, próximo ao brônquio, que é a principal via aérea do corpo humano;
  - Carcinoma de grandes células (indiferenciado)** – pode surgir em qualquer parte do pulmão. Tende a crescer e se espalhar rapidamente, o que dificulta o seu tratamento.
- Câncer de pulmão de pequenas células (CPPC) – de 10% a 15%.

A multiplicação de suas células torna-o mais sensível ao tratamento com quimioterapia e radioterapia, e sua agressividade está relacionada a maiores taxas de recidiva (retorno) da doença. Além disso, cerca de 70% das pessoas com CPPC já possuem metástases no momento do diagnóstico.

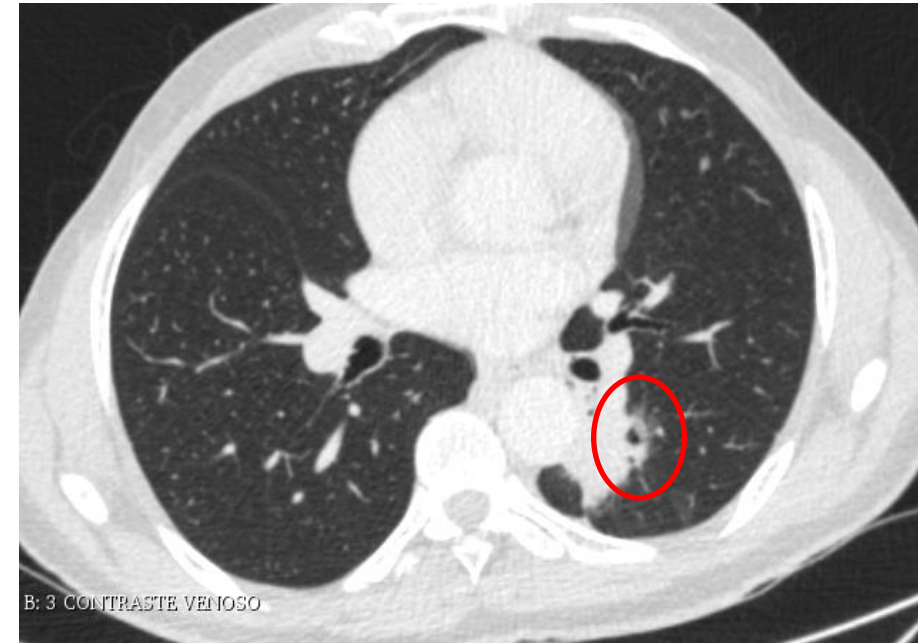


# DIFERENCIAÇÃO DE CAVITAÇÃO E PSEUDOCAVITAÇÃO



## ➤ Cavitação:

- frequentemente associada ao câncer de pulmão de células escamosas
- geralmente indicando mau prognóstico



## ➤ Pseudocavitação:

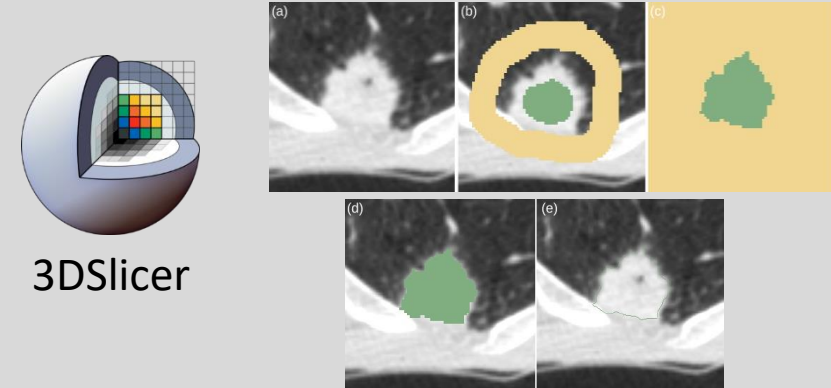
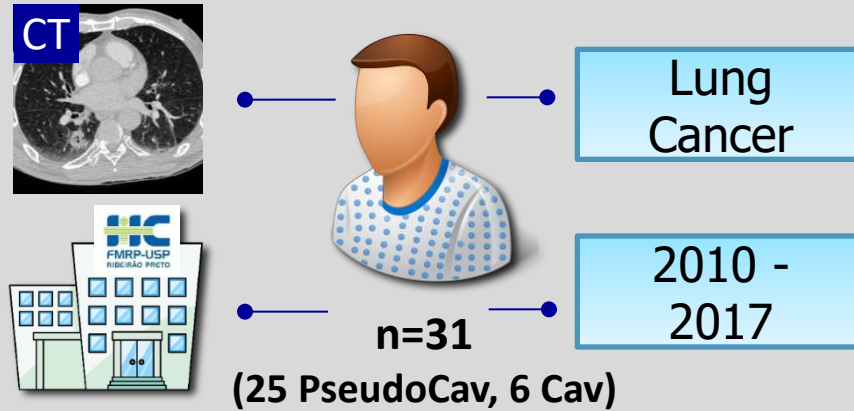
- frequentemente associada a adenocarcinoma pulmonar
- melhor prognóstico com **possibilidade de terapia alvo**

JÚNIOR, José Raniery Ferreira. Arcabouço para Classificação, Recuperação por Conteúdo e Radiômica de Imagens Médicas: uma investigação de biomarcadores quantitativos para o câncer de pulmão. 2019. Tese (Doutorado em Bioengenharia) - Bioengenharia, Universidade de São Paulo, São Carlos, 2019. doi:10.11606/T.82.2020.tde-27022020-113956.

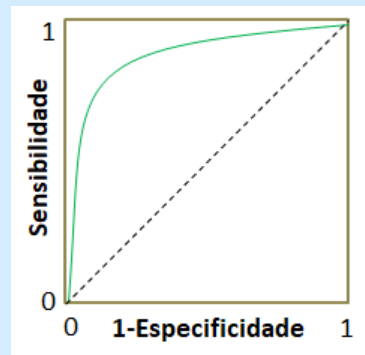
**Prêmio Tese Destaque USP 2020 - Grande Área – Engenharias**

<https://jornal.usp.br/institucional/anunciados-os-vencedores-do-premio-tese-destaque-usp-2020/>

# METHODS



## Mann-Whitney U Test



## 2447 features were extracted

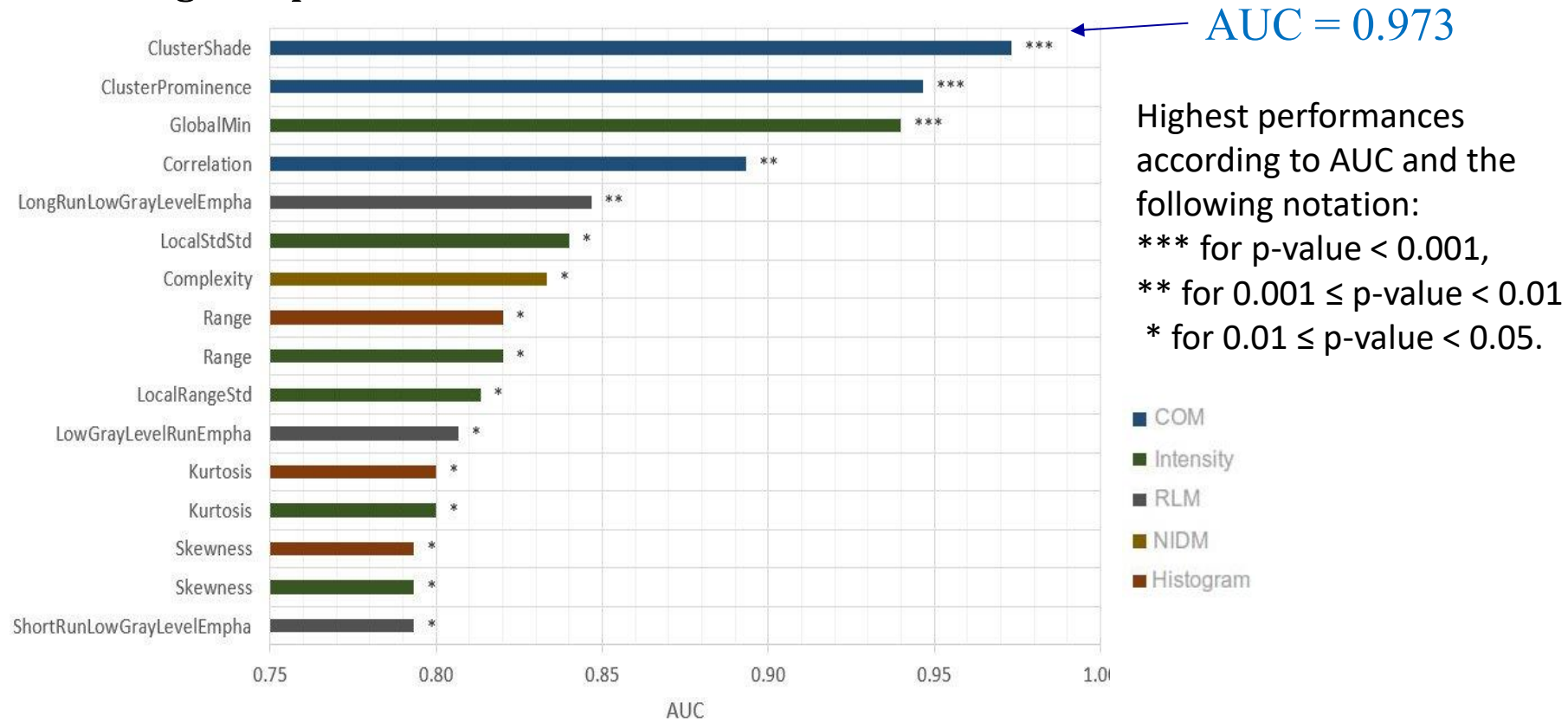
IBEX v1.0



JFeatureLib v1.6.3

## RESULTS

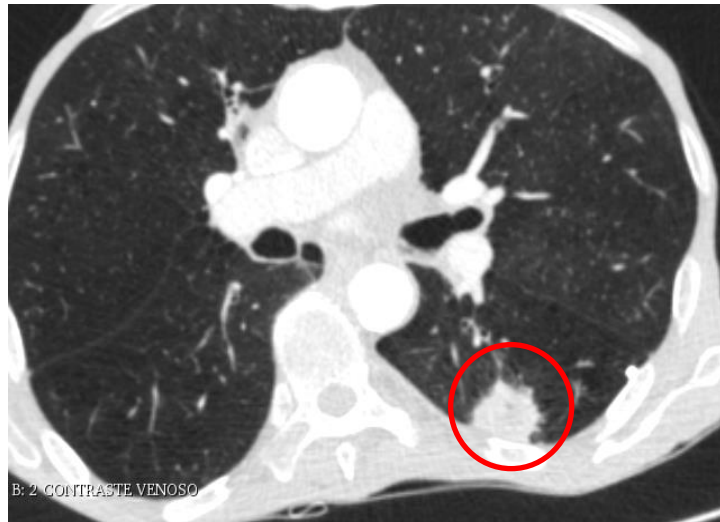
Analysis identified 227 statistically relevant quantitative CT features to distinguish pseudocavitation from cavitation.



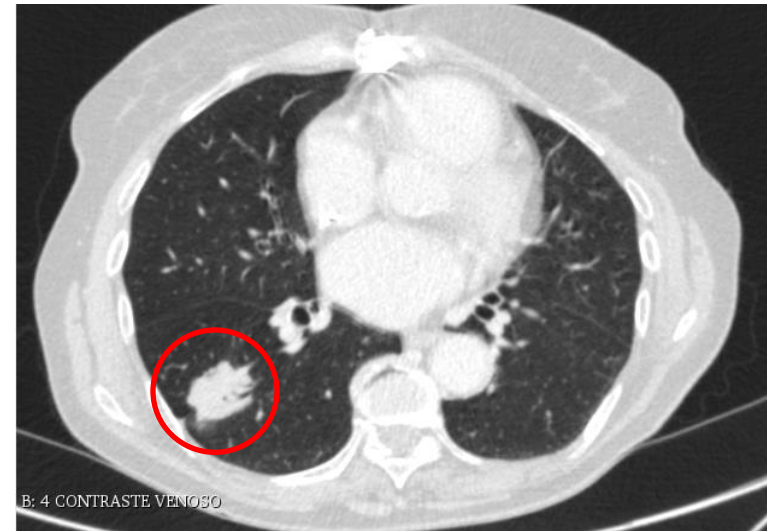
Cooccurrence matrix (COM), Run-length matrix (RLM), Neighborhood intensity-difference matrix (NIDM)



# AVALIAÇÃO PROGNÓSTICA DO PACIENTE (SOBREVIDA)



Adenocarcinoma pulmonar em homem de 68 anos, com ocorrência de óbito.



Adenocarcinoma pulmonar em mulher de 77 anos, sem ocorrência de óbito.

Artigo Original

<http://dx.doi.org/10.1590/0100-3984.2019.0135>

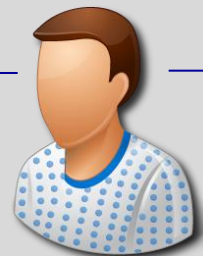
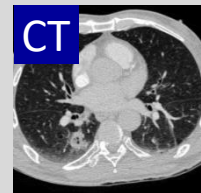
## Análise radiômica do câncer de pulmão para avaliação prognóstica do paciente e da heterogeneidade intratumoral

*Radiomic analysis of lung cancer for the assessment of patient prognosis and intratumor heterogeneity*

José Raniery Ferreira Junior<sup>1,a</sup>, Marcel Koenigkam-Santos<sup>1,b</sup>, Camila Vilas Boas Machado<sup>1,c</sup>, Matheus Calil Faleiros<sup>1,d</sup>, Natália Santana Chiari Correia<sup>1,e</sup>, Federico Enrique Garcia Cipriano<sup>1,f</sup>, Alexandre Todorovic Fabro<sup>1,g</sup>, Paulo Mazzoncini de Azevedo-Marques<sup>1,h</sup>

**Objetivo:** Associar características radiômicas de lesões pulmonares em imagens de tomografia computadorizada com a sobrevida global de pacientes com câncer de pulmão.

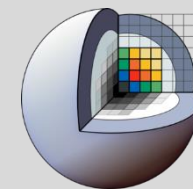
# MÉTODOS



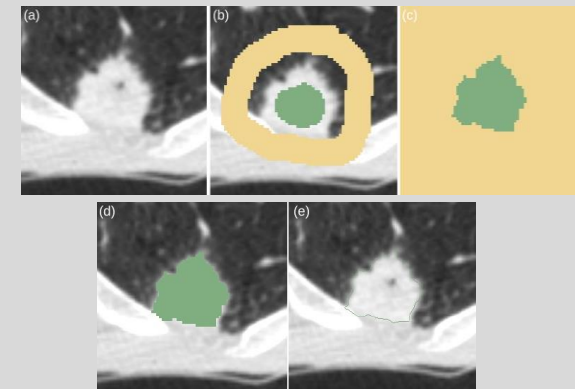
**n=101**

**Câncer de Pulmão**

**2010 - 2017**



**3DSlicer**



## Análise

Avaliação prognóstica(tempo até a morte por qualquer causa): análise de Kaplan-Meier

Estatísticas: Curvas de Kaplan-Meier - Teste log-rank; recursos de prognóstico Radiomics - sobrevivência do pacote R v3.3.3

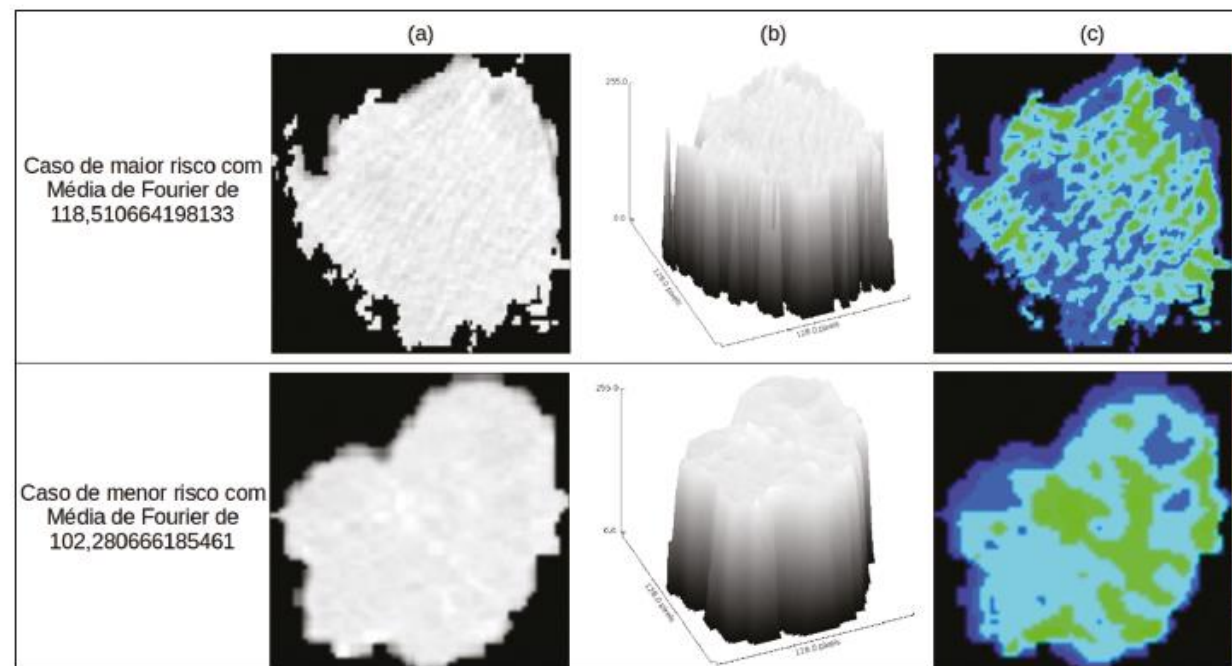
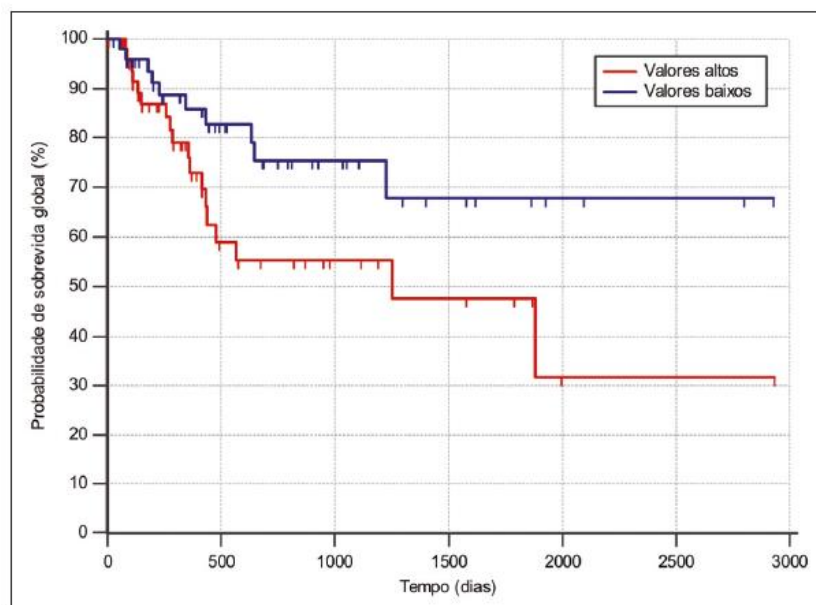
## Foram extraídas 2447 características

**IBEX v1.0**



**JFeatureLib v1.6.3**

## RESULTADOS



**Figura 4.** Quantificação da heterogeneidade intratumoral de lesões estratificadas pela média de Fourier. **a:** Tumor segmentado de imagem de TC em janela de pulmão no plano axial. **b:** Distribuição tridimensional dos níveis de cinza da imagem. **c:** Mapa refletindo a energia local com aplicação de janela de  $5 \times 5$  pixels do tumor.



# Predicting Cardiovascular Events Using Machine Learning (PRE-CARE ML)

- Financiamento: ERA PerMed funded project PRECARE-ML (Reference number: ERAPERMED2021-219) e FAPESP #2021/06137-4
- O projeto PRE-CARE ML tem como objetivo desenvolver modelos para prever o risco de ocorrência de MACE (eventos cardiovasculares adversos maiores), com base em dados extraídos de registros eletrônicos de saúde de pacientes internados, em até 5 anos após a alta hospitalar.
- Coordenador:
  - Peter P. Rainer (Medical University of Graz, University Hospital and Heart Center Graz)
- Pesquisadores Principais:
  - Paulo Mazzoncini de Azevedo-Marques (Ribeirão Preto Medical School at USP)
  - Max Gordon (Karolinska Institutet - Department of Clinical Sciences, Danderyd Hospital)
  - Bert Arnrich (Hasso Plattner Institute for Digital Engineering GmbH)

<https://www.eppermed.eu/funding-projects/projects-results/project-database/pre-care-ml/>

<https://bv.fapesp.br/en/auxilios/110451/predicting-cardiovascular-events-using-machine-learning/>

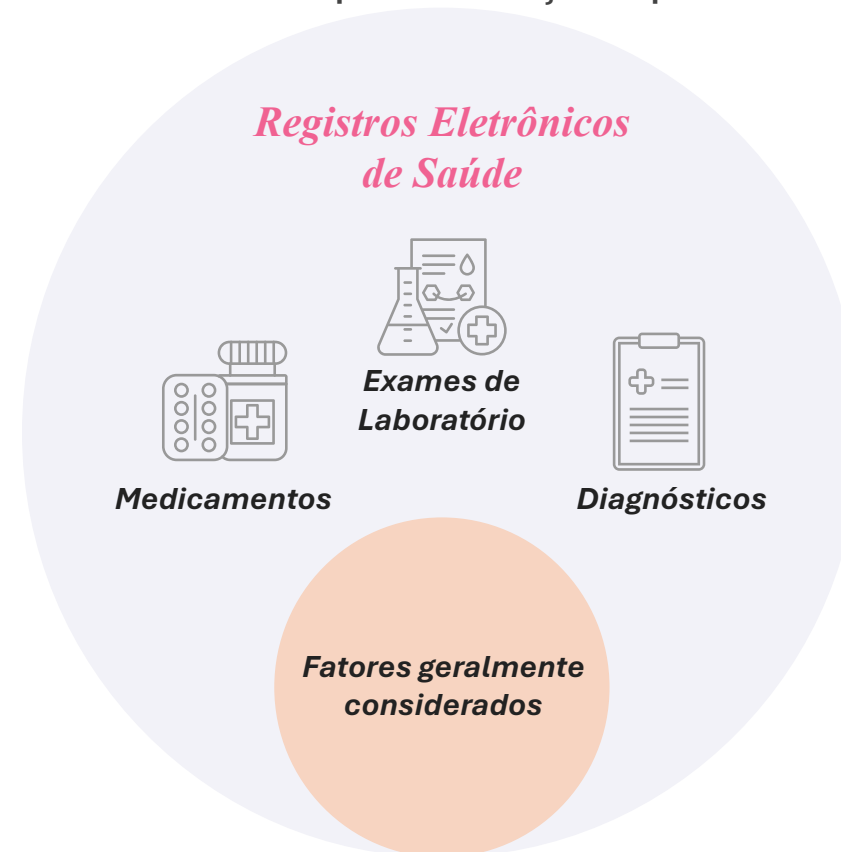
# Predicting Cardiovascular Events Using Machine Learning (PRE-CARE ML)

- A doença cardíaca isquêmica e o acidente vascular cerebral são as principais causas de morbidade no mundo, e seu impacto deverá aumentar, ocupando os dois primeiros lugares até 2050.

*Global Burden of Disease 2021 Study by The Institute for Health Metrics and Evaluation, University Of Washington*

- A detecção precoce de um alto risco potencial de sofrer tais eventos permite ações preventivas.

Fatores geralmente considerados
Idade
Sexo
Raça
Colesterol Total
HDL-C
Pressão Arterial Sistólica
Terapia Antihipertensiva
História de Diabetes Mellitus
Tabagismo
Proteína C-Reativa de Alta Sensibilidade
Histórico Parental de Infarto do Miocárdio



# Representação Esquemática do Projeto

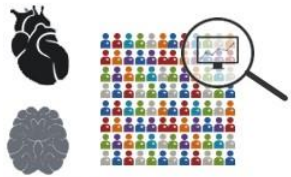
<https://precareml.github.io/>



## PRE-CARE ML

### Predicting Cardiovascular Events Using Machine Learning

#### Aim: Prevention of Cardiovascular Events



Identify patients at high risk



Communicate risk factors and preventive actions

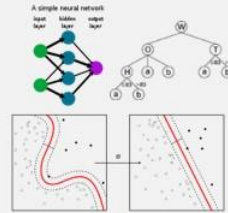


Prevent major adverse cardiovascular events

#### MODEL TRAINING AND VALIDATION



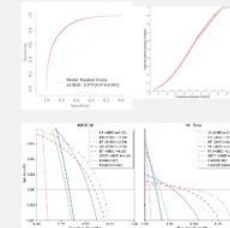
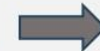
Data pre-processing & harmonization



Machine learning modelling



Federated learning



Evaluation on test data

#### MODEL INTEGRATION & EVALUATION IN HOSPITAL ROUTINE



Hospital admission



Retrieve patient data from system



Automatic risk prediction



Visualize prediction in hospital information system

#### Additional evaluation



Personalized health services for patients



Risk communication strategies



Biases in machine learning models



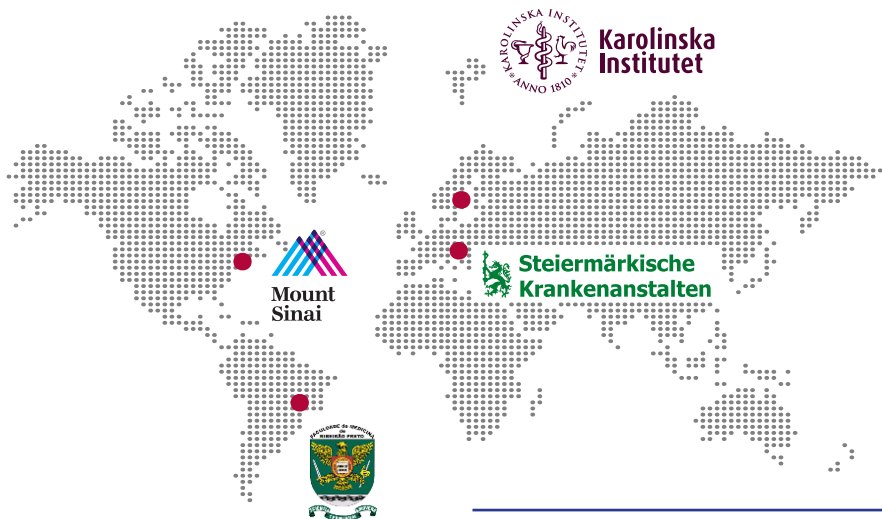
# Dados e Métodos

*Os eventos MACE foram identificados utilizando os códigos de diagnóstico da CID-10 registrados durante as internações hospitalares.*

ICD-10	Descrição	Incluir em MACE?	Justificativa
I20.0	Angina pectoris – Angina instável	Sim	Apenas a angina instável (I20.0) é considerada um evento MACE na maioria das definições.
I21	Infarto agudo do miocárdio	Sim	Parte central de qualquer definição de MACE.
I24	Outras doenças isquêmicas agudas do coração	Sim	Não é claramente definido e inclui condições ambíguas; não é padrão nas definições de MACE.
I63	Infarto cerebral (AVC isquêmico)	Sim	Parte central do MACE.
I71	Aneurisma/dissecção de aorta	Sim	Não faz parte do MACE clássico, mas está associado à aterosclerose. Alguns incluem em MACE de 4 pontos ou definições mais amplas.
I74	Embolia e trombose arteriais	Sim	Possui múltiplas causas possíveis, não sendo específico para aterosclerose.
I20.1	Angina pectoris com espasmo documentado	Não	
I20.8	Outras formas de angina pectoris	Não	
I20.9	Angina pectoris, não especificada	Não	
I46	Parada cardíaca	Não em MACE e NÃO MACE	Pode resultar de várias causas, nem sempre cardiovasculares ou ateroscleróticas.
I64	AVC não especificado como hemorrágico ou isquêmico	Não em MACE e NÃO MACE	Ambíguo; alguns incluem por precaução, outros excluem devido à falta de especificidade.

Feature Name	Description	Time of Observation
SEX	sex: 1 = male / 2 = female	full
AGE	age in years	full
ICD_N_ICD	Number of unique coded ICD-10 codes (3-digit codes only)	3 years (except 10 years for chronic diagnoses)
ICD_W_ICD_<three digit ICD-10>	Weeks since first diagnosis of ICD-10 code (3-digit) (if diagnosed within the last 7 days: value= 1, because 0 means no diagnosis)	3 years (except 10 years for chronic diagnoses)
ICD_N_ICDGRP_<first three digit ICD-10>_<last three digit ICD-10>	Number of unique ICD-10 codes in the ICD10 group e.g. A00 - B99	3 years (except 10 years for chronic diagnoses)
ICD_N_CHRONIC	Count of unique chronic diagnoses (definition chronic: see ICD_chronic_definition)	10 years
ICD_N_CHAP_<ICD-10 chapter>	Unique number of ICD-10 codes in the ICD10 chapter e.g. X (chapter have been transferred from roman numerals)	3 years (except 10 years for chronic diagnoses)
ICD_CHARLSON_WSCORE	Calculated weighted comorbidity score based on charlson comorbidity index (R comorbidity package)	3 years (except 10 years for chronic diagnoses)
LAB_N_<LOINC_CODE>	Number of unique LOINC codes documented in the last 3 years until index date	3 years
LAB_N_ABNORMAL_L_<LOINC_CODE>	Number of low LOINC codes in the last 3 years until index date	3 years
LAB_N_ABNORMAL_0_<LOINC_CODE>	Number of normal LOINC codes in the last 3 years until index date	3 years
LAB_N_ABNORMAL_H_<LOINC_CODE>	Number of elevated LOINC codes in the last 3 years until index date	3 years
LAB_ABNORMAL_<LOINC_CODE>	Latest abnormal category for this LOINC CODE (-1 to 1)	3 years
MEDI_<ATC Level 1>	Any medication of this ATC Level 1 class documented (0 = No 1 = Yes)	3 years
MEDI_<ATC Level 2>	Any medication of this ATC Level 2 class documented (0 = No 1 = Yes)	3 years
MEDI_<ATC Level 3>	Any medication of this ATC Level 3 class documented (0 = No 1 = Yes)	3 years
ADMIN_N_ADMISSION	Number of admissions	5 years
ADMIN_N_OUTPATIENT	Number of outpatient stays	5 years
ADMIN_N_INPATIENT	Number of inpatient stays	5 years

# Dados e Métodos



Lorenzer L, Schrempf M, Jauk S, Kalabakov S, Pierri G, Suzuki KMF, César HV, Kramer D, Mazzoncini de Azevedo-Marques P, Rainer P. **Harmonization of Electronic Medical Records for Federated Learning: Addressing Challenges in International Healthcare Collaborations.** Stud Health Technol Inform. 2025 Apr 24;324:141-147. <https://ebooks.iospress.nl/doi/10.3233/SHTI250175>

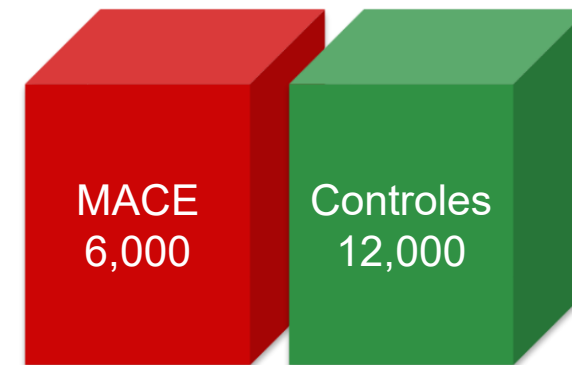
Partner	With MACE	Without MACE	Diagnoses	Structured EHR Modality		
				Laboratory Values	Medication	Procedures
Styrian Hospital Association (Austria)	169,000	263,000	ICD-10	LOINC	Local Standard	Local Standard
Mount Sinai (USA)	7,500	159,000	SNOMED	LOINC	RxNorm	CPT4
Ribeirão Preto Medical School (Brazil)	13,243	84,388	ICD-10	Local Standard	Local Standard	Not Available
Karolinska Institute (Sweden)	/	/	ICD-10	NPU	ATC	ICPM
Chosen Standards			ICD-10	LOINC	ATC	Disregarded

# Dados e Métodos

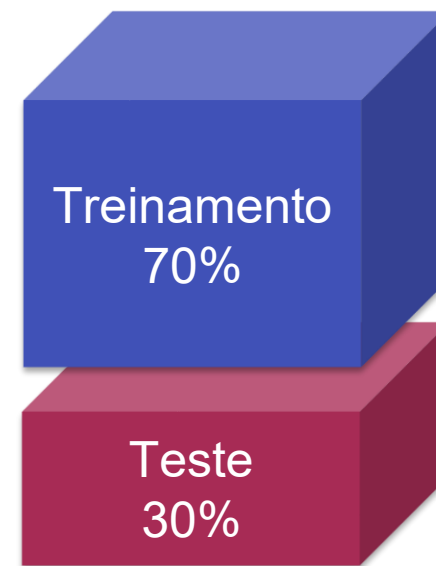
Shimizu GY, Schremppf M, Romão EA, Jauk S, Kramer D, Rainer PP, et al. (2024) **Machine learning-based risk prediction for major adverse cardiovascular events in a Brazilian hospital: Development, external validation, and interpretability.** PLoS ONE 19(10): e0311719. <https://doi.org/10.1371/journal.pone.0311719>

Suzuki, K. M. F., Pierri, G., Shimizu, G. Y., Cesar, H. V., & Azevedo-Marques, P. M. de . (2024). **Brazilian Cohort for Predicting Cardiovascular Events Using Machine Learning (PRE-CARE ML project) (1.0) [Data set].** Zenodo. <https://doi.org/10.5281/zenodo.12752433>

## Dados Desbalanceados



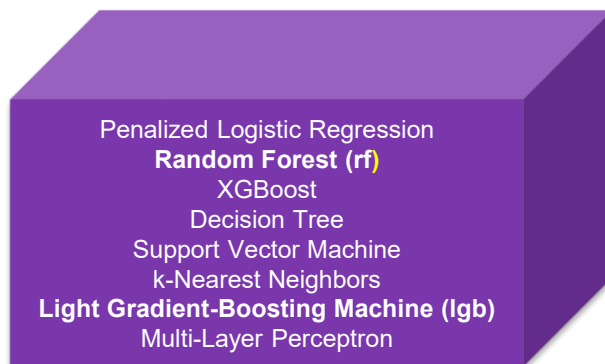
## Divisão Estratificada de Dados



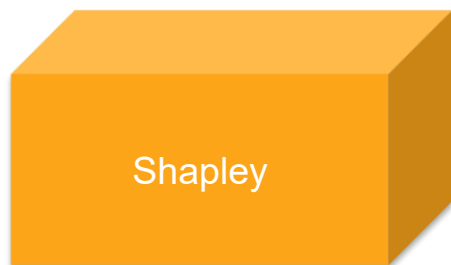
## Geração de Atributos (Python)



## Modelos



## Interpretabilidade

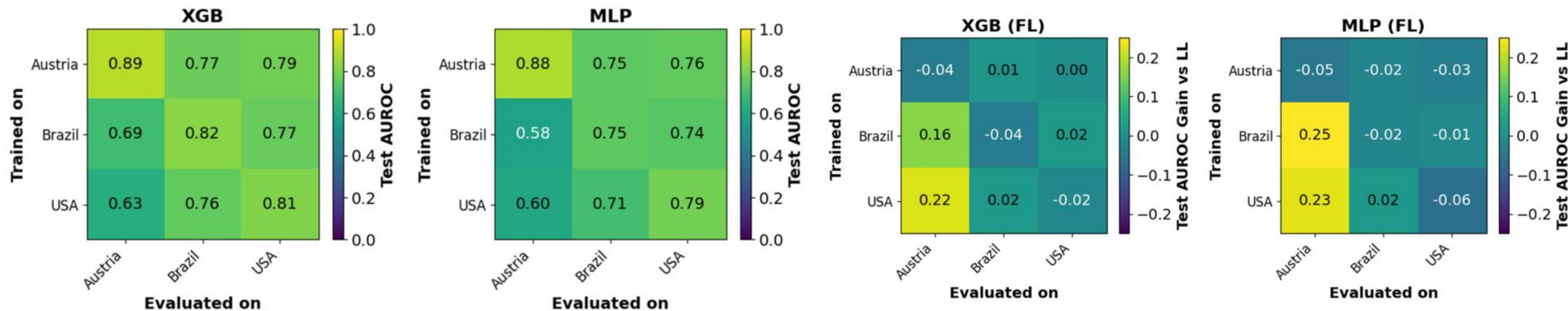




# Resultados

Feature Combination	Trained in	Evaluated in	AUROC * 100
Demo., Admin., Diag.	Austria	Austria	87.04
	USA	USA	83.55
	Austria	USA	69.1
	Brazil	Brazil	74.5
	Austria	Brazil	74.0
Demo., Admin., Diag., Med.	Austria	Austria	88.29
	USA	USA	85.71
	Austria	USA	74.09
	Brazil	Brazil	81.66
Demo., Admin., Diag., Med., Lab.	Brazil	Brazil	91.33
	Austria	Austria	88.96
	USA	USA	87.82
	Austria	USA	74.96

# Resultados Aprendizado Federado (NVIDIA FLARE platform)



Resultados de aprendizagem local para cada combinação de país de treino e teste

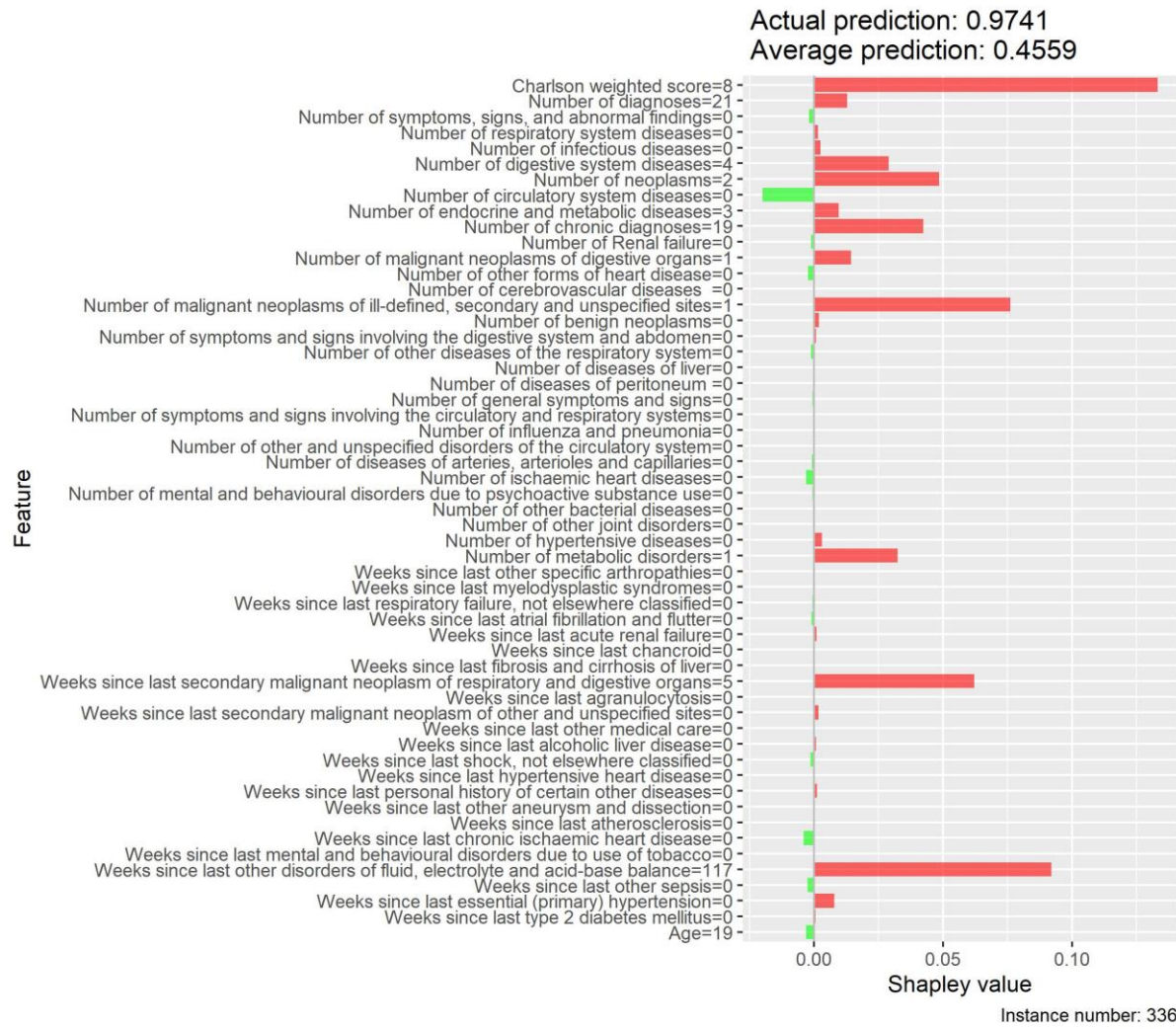
Melhoria na AUROC do modelo FL em relação aos modelos treinados apenas com dados locais.

Model	Austria	Brazil	USA
XGB	0.85	0.78	0.79
MLP	0.83	0.73	0.73

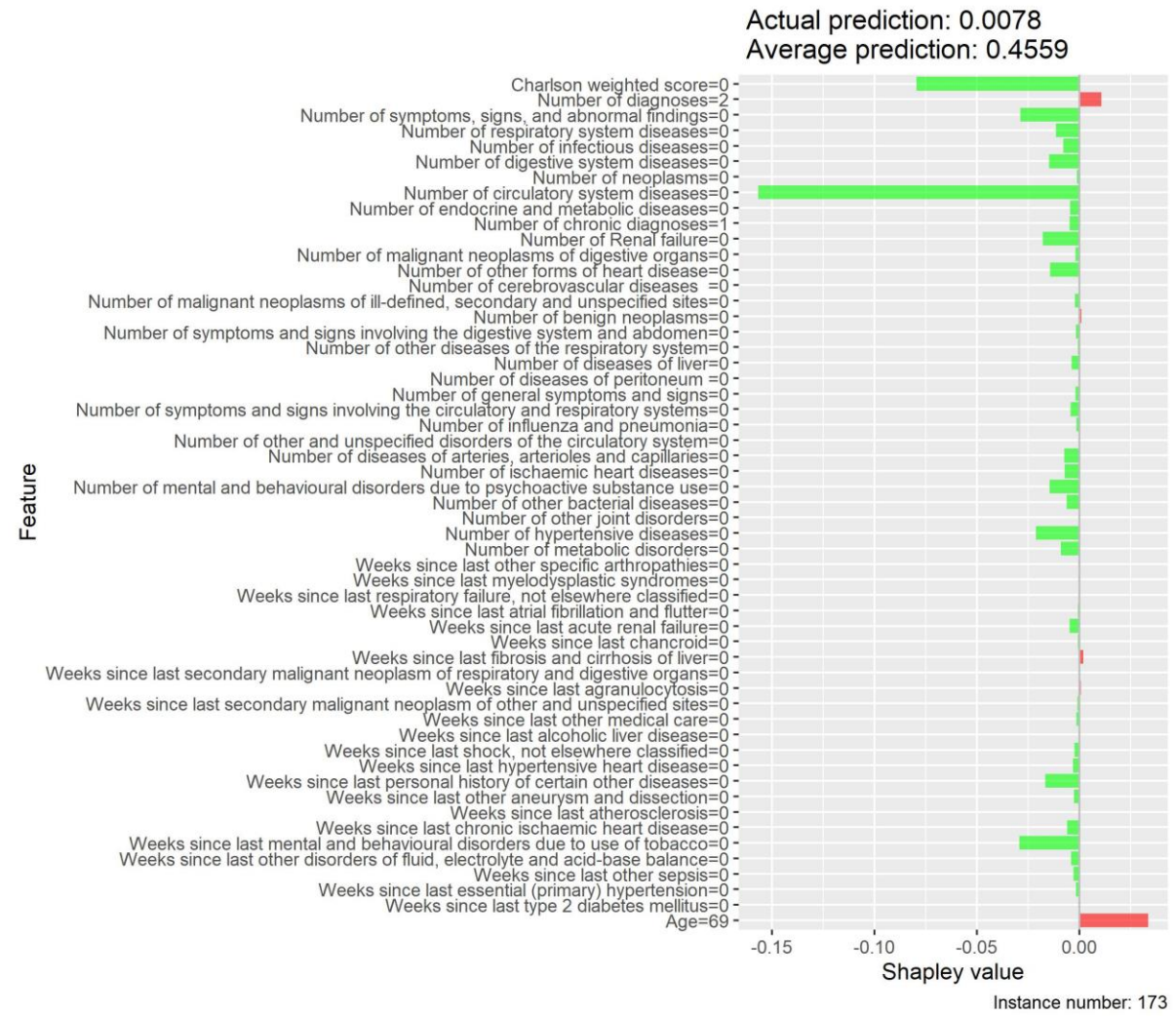
Desempenho AUROC dos modelos FL quando avaliados em cada cliente

Slamanig G, Kalabakov S, Lorenzer L, Schrempf M, Pierri G, Suzuki KMF, Jauk S, Kramer D, Mazzoncini de Azevedo-Marques P, Arnrich B, Rainer P. **Preliminary Investigation of Federated Learning for MACE Prediction from Electronic Medical Records: A Multicontinental Study.** Stud Health Technol Inform. 2026 May 7;335:236-241. doi: 10.3233/SHTI260090. PMID: 42119126. <https://pubmed.ncbi.nlm.nih.gov/42119126/>

# Resultados/Interpretabilidade



(a) Alto risco

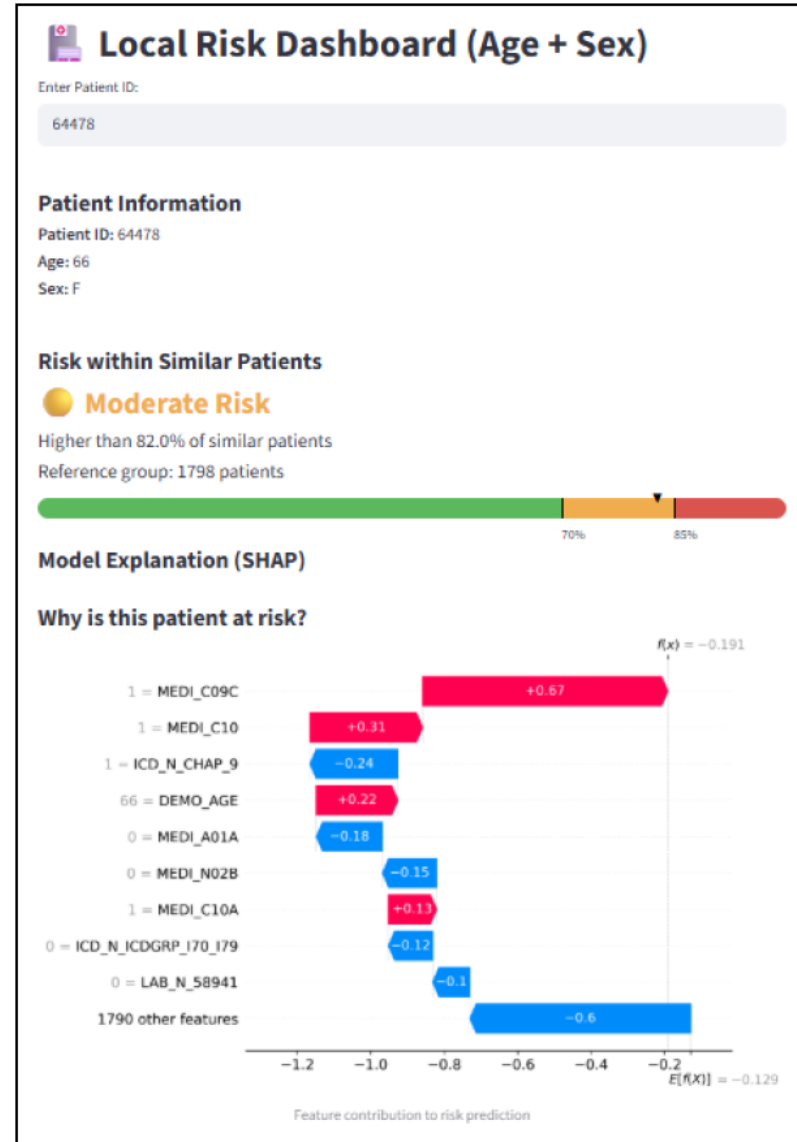
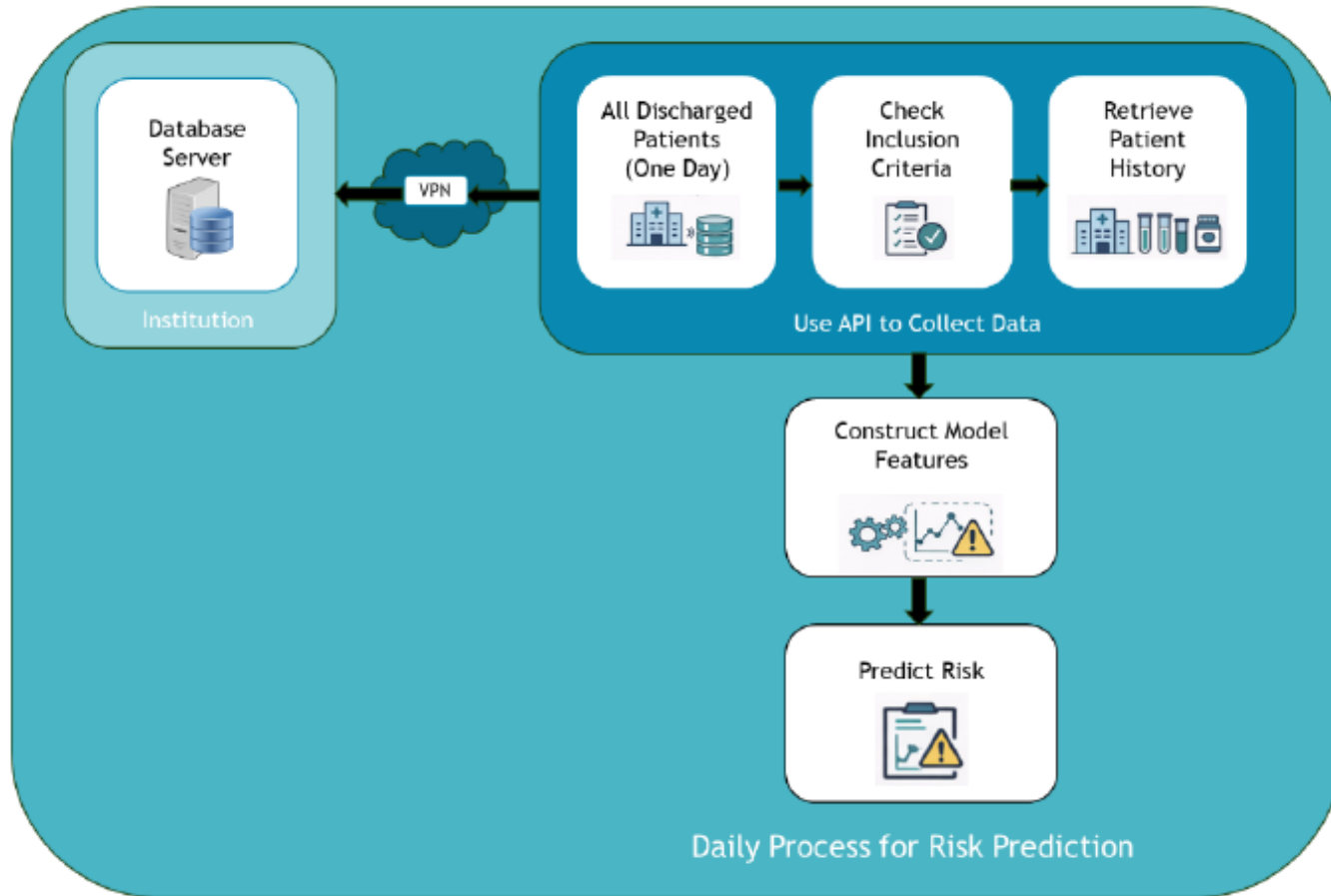


(b) Baixo risco

Interpretabilidade de casos com alto e baixo risco de eventos cardiovasculares adversos maiores (MACE) pelos valores de Shapley.



# Resultados/API/Dashboards





Centro para Inteligência Artificial em Gestão de Saúde  
<https://ciagsaude.icmc.usp.br/pt-BR>



# **Do Diagnóstico à Prevenção: Como a IA Está Antecipando Riscos na Saúde**

Paulo Mazzoncini de Azevedo-Marques E.E Ph.D.

Departamento de Imagens Médicas, Hematologia e Oncologia Clínica

Faculdade de Medicina de Ribeirão Preto - USP

[pmarques@fmrp.usp.br](mailto:pmarques@fmrp.usp.br)