

**UNIVERSIDADE DE SÃO PAULO  
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO**

**LUAN OLIVEIRA BARRETO**

**CAPACIDADE DE PREDIÇÃO DE AGRESSIVIDADE  
DE NEOPLASIAS RENAIIS MALIGNAS UTILIZANDO  
RADIÔMICA EM SEQUÊNCIAS ESPECÍFICAS DE  
RESSONÂNCIA MAGNÉTICA**

RIBEIRÃO PRETO  
2023

**LUAN OLIVEIRA BARRETO**

**CAPACIDADE DE PREDIÇÃO DE AGRESSIVIDADE DE  
NEOPLASIAS RENAIIS MALIGNAS UTILIZANDO  
RADIÔMICA EM SEQUÊNCIAS ESPECÍFICAS DE  
RESSONÂNCIA MAGNÉTICA**

Dissertação apresentada à Faculdade de  
Medicina de Ribeirão Preto da Universidade  
de São Paulo para Titulação de Mestrado  
Profissional.

Área de concentração: Radiologia

Orientador: Prof. Dr. Valdair Francisco  
Muglia

RIBEIRÃO PRETO  
2023

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Barreto, Luan Oliveira

Capacidade de predição de agressividade de neoplasias renais malignas utilizando radiômica em sequências específicas de ressonância magnética. Ribeirão Preto, 2023.

45 p. : 8 il. ; 30cm

Dissertação de Mestrado Profissional, apresentada à Faculdade de Medicina de Ribeirão Preto/USP. Área de Concentração: Radiologia.

Orientador: Prof. Dr. Valdair Francisco Muglia

Palavras-chave: 1) Ressonância Magnética; 2) Inteligência Artificial; 3) PyRadiomics; 4) Sklearn; 5) Random Forest; 6) SVM; 7) Logistic Regression; 8) Neoplasia Renal.

BARRETO, Luan Oliveira.

Capacidade de predição de agressividade de neoplasias renais malignas utilizando radiômica em sequências específicas de ressonância magnética.

Dissertação apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo para titulação de mestrado profissional. Área de concentração: Radiologia

Aprovado em: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

### **BANCA EXAMINADORA**

Prof. Dr. \_\_\_\_\_

Instituição \_\_\_\_\_

Julgamento \_\_\_\_\_ Assinatura \_\_\_\_\_

Prof. Dr. \_\_\_\_\_

Instituição \_\_\_\_\_

Julgamento \_\_\_\_\_ Assinatura \_\_\_\_\_

Prof. Dr. \_\_\_\_\_

Instituição \_\_\_\_\_

Julgamento \_\_\_\_\_ Assinatura \_\_\_\_\_

## **AGRADECIMENTOS**

Aos meus pais e irmãos, pelo amor e apoio incondicional todos os dias e em todas as fases da minha trajetória pessoal e profissional.

Ao meu orientador, Prof. Dr. Valdair Muglia, pela amizade, confiança em meu potencial e ensinamentos neste trabalho e na vida.

Ao meu colega Júlio Nather, por ajudar a concretizar ideias no computador.

Aos colegas Carlos Mikio e Rafael Barros, que me ajudaram a fazer esse projeto sair do papel.

Ao Prof. Dr. Jorge Elias, Prof. Dr. Marcel Koenigkam e Prof. Dr. Danilo Wada pela amizade e confiança em meu potencial.

Por fim, a todos os amigos que fiz em Ribeirão Preto, que contribuíram de alguma forma nesse projeto.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

## **NÚCLEO DE PESQUISA/SETOR**

Centro de Ciências das Imagens e Física Médica da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo - USP.

### **LOCAL DE EXECUÇÃO**

Unidade Campus do Hospital das Clínicas de Ribeirão Preto - Universidade de São Paulo - USP.

## **PESQUISADORES RESPONSÁVEIS**

Dr. Luan Oliveira Barreto

Pós-graduando em Radiologia e Diagnóstico por Imagem da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo - USP.

Dr. Júlio César Nather Júnior

Doutorando em Radiologia e Diagnóstico por Imagem da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo - USP.

Dr. Carlos Mikio Tanno Filho

Médico Adido do Setor de Radiologia e Diagnóstico por Imagem da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo - USP.

Dr. Rafael Vasconcelos Barros

Pós-graduando em Radiologia e Diagnóstico por Imagem da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo - USP.

Prof. Dr. Jorge Elias Junior

Professor Doutor do Departamento de Imagens Médicas, Oncologia e Hematologia da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo - USP.

Prof. Dr. Valdair Francisco Muglia

Professor Doutor do Departamento de Imagens Médicas, Oncologia e Hematologia da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo - USP.

## LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

<b>ADC</b>	<i>Apparent diffusion coefficient</i> - Coeficiente de difusão aparente
<b>AUC</b>	<i>Area under the curve</i> - Área sob a curva
<b>CCR</b>	Carcinoma de células renais
<b>CCRcc</b>	Carcinoma de células renais subtipo células claras
<b>CCRp</b>	Carcinoma de células renais subtipo papilífero
<b>DICOM</b>	<i>Digital imaging and communications in Medicine</i> - Imagens digitais e comunicações em Medicina
<b>DWI</b>	<i>Diffusion-weighted Imaging</i> - Sequência com restrição à difusão
<b>GLCM</b>	<i>Grey level cooccurrence matrix</i> - Matriz de coocorrência em nível de cinza
<b>GLSZM</b>	<i>Grey level size zone matrix</i> - Matriz de zona de tamanho em nível de cinza
<b>GLDM</b>	<i>Grey level dependence matrix</i> - Matriz de dependência de nível de cinza
<b>GLRLM</b>	<i>Grey level run length matrix</i> - Matriz de comprimento de execução em nível de cinza
<b>NGTDM</b>	<i>Neighboring gray tone difference matrix</i> - Matriz de diferença entre tons de cinza vizinhos
<b>HCFMRP - USP</b>	Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo
<b>IA</b>	Inteligência artificial
<b>ISUP</b>	<i>International Society of Urologic Pathologists</i> – Sociedade Internacional de Patologistas Urológicos
<b>PET</b>	<i>Prostate Imaging Reporting &amp; Data System</i> - Sistema de dados e relatórios de imagem da próstata
<b>RIS</b>	<i>Radiological information system</i> - Sistema de informações radiológicas
<b>RM</b>	Ressonância magnética
<b>ROC</b>	<i>Receiver operator characteristic</i>
<b>VOI</b>	<i>Volume of interest</i> - Volume de interesse
<b>SIH</b>	Sistema de informações hospitalares
<b>SVM</b>	<i>Support Vector Machine</i> - Máquina de vetores de suporte
<b>PCA</b>	<i>Principal component analysis</i> - Análise do componente principal

## RESUMO

Barreto, L. O. **Capacidade de predição de agressividade de neoplasias renais malignas utilizando radiômica em sequências específicas de ressonância magnética.** 2023. Dissertação (Mestrado Profissional). Faculdade de Medicina de Ribeirão Preto, Universidade da São Paulo, Ribeirão Preto, 2023.

**Objetivo:** Avaliar a capacidade de sequências específicas de RM em prever a gravidade anatomopatológica de neoplasias renais malignas, utilizando o desempenho da radiômica, baseada em ferramentas de processamento por inteligência artificial com aprendizado de máquina (*machine learning*), treinada com base em exames de pacientes com confirmação histopatológica de câncer de rim. **Materiais e métodos:** Segmentamos, manualmente, 42 lesões renais de exames de RM, realizadas em nossa instituição. Os dados foram adquiridos nas fases T2, ADC e Portal, através do software 3D *Slicer*. Em seguida, foram extraídos atributos (*features*) das imagens usando a biblioteca *PyRadiomics*. Após a extração dos atributos, realizou-se uma seleção de atributos utilizando a biblioteca *Sklearn*. Foram testados diferentes modelos de seleção, como *Recursive Feature Elimination* (RFE) e *SelectKBest*. Os modelos de aprendizado de máquina testados foram *Random Forest*, *Sector Vector Machine* (SVM) e *Logistic Regression*. O treino foi feito com a porção de treino dos casos, e a validação dos modelos com a porção de teste. Após a validação dos modelos, estudamos a acurácia, a precisão e a área sob a curva ROC (AUC-ROC) para avaliar o desempenho dos modelos. **Resultados:** no ADC, o SVM, *Random Forest* e *Logistic Regression* alcançaram acurácia de 0.80, 0.68 e 0.78, precisão de 0.65, 0.46 e 0.59 e AUC-ROC de 0.73, 0.57 e 0.73. No Portal, o SVM, *Random Forest* e *Logistic Regression* alcançaram acurácia de 0.82, 0.83 e 0.78, precisão de 0.69, 0.62 e 0.64 e AUC-ROC de 1.00, 0.67 e 0.82. **Discussão:** os resultados deste estudo demonstram que a radiômica pode ser usada para diferenciar entre tumores renais de baixo e alto grau usando dados de ressonância magnética com boa precisão. O SVM e *Logistic Regression* obtiveram bons resultados, com SVM alcançando a maior precisão e o *Random Forest* não obteve bons resultados. Novos estudos com amostras maiores são necessários para confirmar esses achados.

**Palavras chave:** Ressonância Magnética; Inteligência Artificial; *PyRadiomics*; *Sklearn*; *Random Forest*; SVM; *Logistic Regression*; Neoplasia Renal.

## ABSTRACT

Barreto, L. O. **Ability to predict the anatomopathological severity of malignant renal neoplasms using radiomics in specific sequences of Magnetic Resonance.** 2023. Dissertation (Professional Master's Degree). Faculty of Medicine of Ribeirão Preto, University of São Paulo, Ribeirão Preto, 2023.

**Objective:** To evaluate the ability of specific MRI sequences to predict the anatomopathological severity of malignant renal neoplasms, using the performance of radiomics, based on processing tools by artificial intelligence with machine learning, trained on the basis of patient exams with histopathological confirmation of kidney cancer. **Materials and methods:** We manually segmented 42 renal lesions from MRI scans performed at our institution. Data were acquired in phases T2, ADC and Portal, through the 3D Slicer software. Then, features of the images were extracted using the PyRadiomics library. After extracting the attributes, a selection of attributes was performed using the Sklearn library. Different selection models were tested, such as Recursive Feature Elimination (RFE) and SelectKBest. The tested machine learning models were Random Forest, Sector Vector Machine (SVM) and Logistic Regression. Training was done with the training portion of the cases, and validation of the models with the test portion. After validating the models, we studied the accuracy, precision and area under the ROC curve (AUC-ROC) to evaluate the performance of the models. **Results:** in the ADC, the SVM, Random Forest and Logistic Regression achieved accuracy of 0.80, 0.68 and 0.78, precision of 0.65, 0.46 and 0.59 and AUC-ROC of 0.73, 0.57 and 0.73. In Portal, SVM, Random Forest and Logistic Regression reached accuracy of 0.82, 0.83 and 0.78, precision of 0.69, 0.62 and 0.64 and AUC-ROC of 1.00, 0.67 and 0.82. **Discussion:** The results of this study demonstrate that radiomics can be used to differentiate between low and high grade renal tumors using MRI data with good accuracy. SVM and Logistic Regression achieved good results, with SVM achieving the highest accuracy and Random Forest not achieving good results. New studies with larger samples are needed to confirm these findings.

**Keywords:** Magnetic Resonance; kidney neoplasm; artificial intelligence; PyRadiomics; Sklearn; Random Forest; SVM; Logistic Regression; kidney neoplasm

## SUMÁRIO

1	INTRODUÇÃO.....	12
2	OBJETIVOS.....	16
	2.1 Objetivo Principal:.....	16
	2.2 Objetivo Secundário .....	16
3	MATERIAIS E MÉTODOS .....	17
	3.1 Tipo de estudo.....	17
	3.2 Local do estudo.....	17
	3.3 Coleta de dados.....	17
	3.4 Comitê de Ética.....	17
	3.5 Critérios de inclusão .....	18
	3.6 Critérios de exclusão.....	18
	3.7 Protocolo de Ressonância Magnética .....	19
	3.8 Segmentação na ressonância magnética .....	20
	3.9 Randomização dos pacientes .....	21
	3.10 Extração de recursos .....	22
	3.11 Análise e processamento dos dados.....	23
	3.11.1 Análise exploratória dos dados.....	23
	3.11.2 Redução de dimensionalidade .....	24
	3.11.3 Visualização de dendrogramas e hierarquias de agrupamento.....	24
	3.11.4 Preparação dos dados para treinamento e validação .....	24
	3.11.5 Normalização e padronização dos dados.....	24
	3.11.6 Treinamento e avaliação dos modelos de classificação.....	25
	3.11.7 Validação cruzada.....	25
	3.11.8 Comparação e seleção do melhor modelo .....	26
	3.11.9 Implementação e monitoramento do modelo .....	26
	3.12 Análise estatística: .....	26
4	RESULTADOS .....	27
5	DISCUSSÃO.....	35
6	CONCLUSÃO.....	39
7	REFERÊNCIAS BIBLIOGRÁFICAS-.....	40

## LISTA DE FIGURAS

Figura 1 - Fluxograma mostrando número de exames iniciais, exclusões e o estudo de corte final. ....	19
Figura 2 - Segmentação de lesão renal visualizada na RM no 3D Slicer, com imagens em axial antes e após segmentação, com reconstruções sagital e coronal.....	21
Figura 3 - Mapa ilustrativo do funcionamento do PyRadiomics.....	23
Figura 4 - Distribuição dos tumores por idade e sexo. ....	28
Figura 5 - Distribuição por idade separada para cada sexo, por grupo de agressividade anatomopatológica dos tumores. ....	28
Figura 6 - Comparação dos modelos classificadores com <i>cross validation</i> para o grupo ADC.....	30
Figura 7 - Comparação dos modelos classificadores com <i>cross validation</i> para o grupo Portal.....	32
Figura 8 - Comparação dos modelos classificadores com <i>cross validation</i> para o grupo T2.....	34

## LISTA DE TABELAS

Tabela 1 - Distribuição dos pacientes por sexo .....	27
Tabela 2 - Distribuição dos pacientes por idade.....	27
Tabela 3 - Acurácia, precisão e área sob a curva ROC dos diferentes classificadores utilizados para o grupo ADC.....	30
Tabela 4 - Acurácia, precisão e área sob a curva ROC dos diferentes classificadores utilizados para o grupo Portal .....	31
Tabela 5 - Acurácia, precisão e área sob a curva ROC dos diferentes classificadores utilizados para o grupo T2.....	33

# 1 INTRODUÇÃO

O carcinoma de células renais (CCR) constitui aproximadamente 3,8% dos novos cânceres, é o tipo de câncer renal mais comum e é responsável por cerca de 2% de todas as mortes oncológicas em todo o mundo [1]. O carcinoma de células renais de células claras (CCRcc) é o subtipo mais comum, representando até aproximadamente 80% dos CCR. Entre os diferentes subtipos histológicos, o CCRcc mostrou ter um prognóstico pior do que outros subtipos histológicos (papilar e cromóforo) [2].

O passo mais importante para um médico durante o diagnóstico e tratamento do câncer é o estadiamento e classificação do tumor, pois representa os mais importantes preditores prognósticos da taxa de sobrevivência em 5 anos, que é de 93% para tumor localizado e diminui para 69% e 12% se houver acometimento para os linfonodos e partes distantes do corpo, respectivamente [3].

A classificação do tumor é uma descrição da diferenciação das células do tecido tumoral em relação às células do tecido normal. É um indicador da rapidez com que se espera que um tumor cresça e se espalhe. Atualmente, o padrão-ouro para classificação e graduação do CCR é a avaliação histopatológica do tecido obtido através de biópsia ou cirurgia. Entretanto, biópsia e cirurgia são procedimentos invasivos que carregam riscos e podem não ser viável para muitos pacientes.

Muitos sistemas de classificação foram propostos para neoplasia de células renais, no entanto, o sistema de Fuhrman tradicional de 4 camadas foi o sistema de classificação mais frequente em CCRcc e CCR papilar [4]. O sistema de graduação de Fuhrman é amplamente aceito e baseia-se na avaliação das seguintes características do núcleo celular: tamanho nuclear, forma nuclear e proeminência nucleolar. Devido à alta heterogeneidade espacial e temporal do CCRcc, a biópsia não pode refletir o grau de Fuhrman de todo o tumor [5], portanto, a avaliação pré-operatória do CCR usando procedimento não invasivo é desejável.

Além disso, há o surgimento de consensos e novas classificações quanto aos tipos de tumores renais. Em 2012, a Conferência de Consenso da *International Society of Urological Pathology* (ISUP) fez recomendações sobre a classificação, fatores prognósticos, estadiamento e avaliação molecular e imuno-histoquímica dos tumores renais do adulto. Houve consenso que cinco entidades devem ser reconhecidas como novos tumores: CCR tubulocístico, CCR associado a doença cística adquirida, CCR papilífero de células claras,

CCR de translocação associada ao fator de transcrição da microftalmia e a síndrome de leiomiomatose hereditária com CCR. Houve ainda a inclusão de três tipos adicionais consideradas entidades novas e emergentes [6]. Ainda, foi proposto um novo sistema de gradação pela ISUP, levando também em consideração a avaliação nucleolar em diferentes magnificações [6].

Diante dessa realidade de heterogeneidade da neoplasia renal e do crescimento gradativo não só da qualidade e quantidade dos métodos de imagem, da quantidade de exames de imagem realizados nos serviços de saúde e da detecção precoce de nódulos renais, além do crescimento exponencial das capacidades tecnológicas vigentes, a IA ganha cada vez mais evidência na avaliação por métodos de imagem nas mais diversas situações e patologias.

Por definição, a IA representa a capacidade das máquinas de mimetizar a função cognitiva dos seres humanos para identificar e resolver problemas [7], podendo ser utilizada de inúmeras maneiras no âmbito clínico e diagnóstico. Na área da Radiologia, a IA surgiu como ferramenta capaz de ajudar tanto na otimização da produtividade quanto na melhora da acurácia diagnóstica e predição de desfechos [7, 8, 9, 10, 11, 12, 16].

O processo do uso de IA na Radiologia passa por três definições: aprendizado de máquina, *deep learning* e *data augmentation*. Aprendizado de máquina representa a fase em que radiologistas especialistas amplamente treinados ensinam as características de imagem suspeitas e clinicamente significativas, de modo a aumentar a reprodutibilidade e eficiência no diagnóstico pela máquina [7]. Esta ferramenta é aprimorada com *deep learning*, que é um subconjunto de aprendizados da máquina e, por definição, é a fase que a máquina vai mudando e melhorando a sua capacidade de interpretação à medida que substratos mais numerosos e complexos são adicionados como treinamento, desta vez sem a necessidade de definição prévia por especialistas [8]. Ainda, considerando que uma mesma patologia pode apresentar características de imagem divergentes e mesmo assim convergir à mesma interpretação final, a máquina deve ser ensinada sobre todos os possíveis padrões para uma mesma doença, o que define o conceito de *data augmentation* [7].

A radiômica, que significa os processos de extração de alto rendimento de características quantitativas que resultam na conversão de imagens em dados lavráveis e a posterior análise desses dados para suporte à decisão, incluindo correlação com informações clínicas, diagnósticas e prognósticas, tem o uso crescido exponencialmente na área médica. [8]. Em outras palavras, refere-se à quantificação abrangente de fenótipos tumorais para descobrir características da doença que não podem ser explicadas a olho nu [9].

O processo da IA e radiômica na Radiologia começa com a construção de banco de dados de imagens de pacientes sobre os quais a evolução, desfecho clínico e prognóstico são conhecidos. Quanto maior o banco de dados, mais sensível e específica se torna a capacidade de acerto da máquina, resultando em maior acurácia. Para reduzir erros de interpretação em consequência de artefatos de imagem ou de técnica não padronizada, o ideal é que o computador seja treinado com exames realizados em um mesmo aparelho, sob os mesmos protocolos [10]. Após, através de análise matemática da distribuição espacial da interrelação de pixels e intensidades de sinal, a radiômica pode quantificar essa informação textural oculta em um tecido [11]. Essa informação pode ser cruzada com informações já conhecidas dos pacientes, como desfechos clínicos e resultados de outros exames, como análise anatomopatológica. No campo da oncologia, especialmente, vem crescendo exponencialmente nos últimos anos o número de pesquisas abordando a aplicação da radiômica nos dados de imagens médicas para classificar de forma pré-operatória o tipo ou estágio clínico, ou mesmo prever o prognóstico e desfecho clínico [12].

No cenário clínico e radiológico atual da avaliação por imagem tradicional no CCRcc existem um proeminente obstáculo: a precisão limitada em avaliar dignamente os pequenos tumores renais. Por exemplo, até 30% dos nódulos ressecados após nefrectomia parcial ou total apresentam histologia benigna na avaliação histopatológica [13]. Além disto, mesmo podendo na maioria das vezes ser tumores de baixa agressividade, podem eventualmente representar uma detecção precoce de tumores de alta agressividade.

A ressonância magnética (RM) é comumente a modalidade de escolha para avaliação de tumores renais e tem a vantagem de oferecer informação anatômica detalhada e bom contraste entre os diferentes tipos de tecido. Uma aplicação potencial da radiômica usando RM é na diferenciação de tumores renais de baixo e alto grau. Investigações anteriores mostraram que a análise de textura por TC pode ser usada para diferenciar angiomiolipoma renal pobre em gordura do CCR e assinaturas radiômicas de PET/RM foram associadas à vascularização do tumor e à expressão do fator de crescimento endotelial vascular no CCRcc [14,15]. Realizar uma graduação acurada do CCR é fundamental para determinar o curso de tratamento ideal, incluindo tratamentos mais agressivos como cirurgia e quimioterapia, além de poder prever o desfecho clínico. Outros estudos anteriores mostraram que a assinatura radiômica é valiosa para diferenciar tumores CCRcc de alto e baixo grau [16], o que pode ajudar a prever o prognóstico.

Considerando a importância da classificação histopatológica do carcinoma de células renais como preditor prognóstico e a necessidade de uma avaliação pré-operatória usando

procedimento não-invasivo, faz-se necessário estudar ferramentas de exames de imagem que possam prever o grau de diferenciação tumoral da classificação histopatológica.

## **2 OBJETIVOS**

### **2.1 Objetivo Principal:**

O objetivo deste estudo é investigar a capacidade de características radiômicas baseadas em RM, notadamente ponderadas em T2, ADC e sequências pós-contraste, para determinar o grau de Fuhrman de CCR.

### **2.2 Objetivo Secundário**

Identificar qual modelo de aprendizado de máquina tem maior precisão diagnóstica na análise proposta.

## **3 MATERIAIS E MÉTODOS**

### **3.1 Tipo de estudo**

Trata-se de um estudo retrospectivo, observacional, realizado em um único centro.

### **3.2 Local do estudo**

O estudo foi realizado no Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (HCFMRP-USP).

### **3.3 Coleta de dados**

O Sistema de Informações Radiológicas (RIS) e o Sistema de Informações Hospitalares (SIH) de nossa instituição foram as fontes de pesquisa para o presente estudo. De forma retrospectiva, coletamos dados sobre pacientes submetidos à RM de abdome, no período de maio de 2020 a julho de 2021, e que possuem resultados histopatológicos por meio de biópsia direcionada ou de nefrectomia parcial ou total.

### **3.4 Comitê de Ética**

O estudo foi aprovado pelo Comitê de Ética e Pesquisa (CEP) do HCFMRP-USP com o título: “Estudo sobre a capacidade de predição de gravidade anatomopatológica de neoplasias renais malignas utilizando radiômica em sequências específicas de Ressonância Magnética”, com dispensa do Termo de Consentimento Livre e Esclarecido (TCLE), devido à natureza retrospectiva.

### **3.5 Critérios de inclusão**

Foram incluídos no estudo pacientes diagnosticados com carcinoma de células renais através de estudo histopatológico por análise da peça cirúrgica ou por biópsia tecidual com agulha grossa e que fizeram RM abdominal no período estipulado.

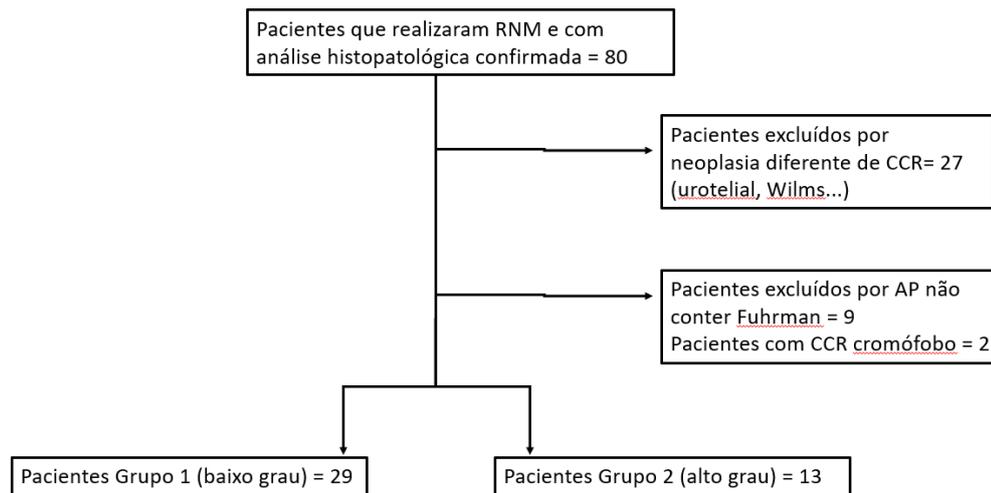
### **3.6 Critérios de exclusão**

Foram excluídos do estudo pacientes nas seguintes situações:

- a) Pacientes com neoplasia renal com diagnóstico histopatológico que não represente um subtipo de carcinoma de células renais
- b) Pacientes com exame de imagem apresentando qualidade inadequada em todas as sequências (T2, ADC e pós-contraste).
- c) Pacientes que a análise histopatológica por biópsia não foi informado grau de Fuhrman.

Desta forma, 80 pacientes que realizaram RM no período estipulado e com análise histopatológica confirmada de neoplasia renal foram incluídos inicialmente no estudo. Houve exclusão de 27 pacientes por não representar subtipo de CCR, como carcinoma urotelial, tumor de Wilms e tumor rabdoide. Após, os resultados das biópsias foram analisados minuciosamente e houve exclusão de mais 9 pacientes devido às biópsias não informarem adequadamente a classificação de Fuhrman e mais 2 pacientes foram excluídos por representar CCR subtipo cromóforo, o qual não é recomendado uso da classificação de Fuhrman.

Após, os pacientes foram divididos em dois grupos baseados na classificação de Fuhrman, sendo o grupo 1 pacientes classificados como grau 1 e 2 de Fuhrman e grupo 2 pacientes classificados como grau 3 e 4 de Fuhrman.



(Fonte: Arquivo próprio)

Figura 1 - Fluxograma mostrando número de exames iniciais, exclusões e o estudo de coorte final.

### 3.7 Protocolo de Ressonância Magnética

Todos os exames foram realizados com um aparelho de ressonância magnética de alta definição, 1.5 Tesla. O protocolo utilizado foi o protocolo padrão de exames de abdome superior utilizado no serviço.

O protocolo inclui sequências ponderadas em T1, ponderadas em T2, imagens ponderadas em difusão (DWI), coeficiente de difusão aparente (ADC) e dinâmicas com contraste. Essas sequências são selecionadas para proporcionar uma melhor visualização e diferenciação das estruturas anatômicas e patológicas da abdome, auxiliando assim na identificação e caracterização de cânceres, incluindo as neoplasias renais.

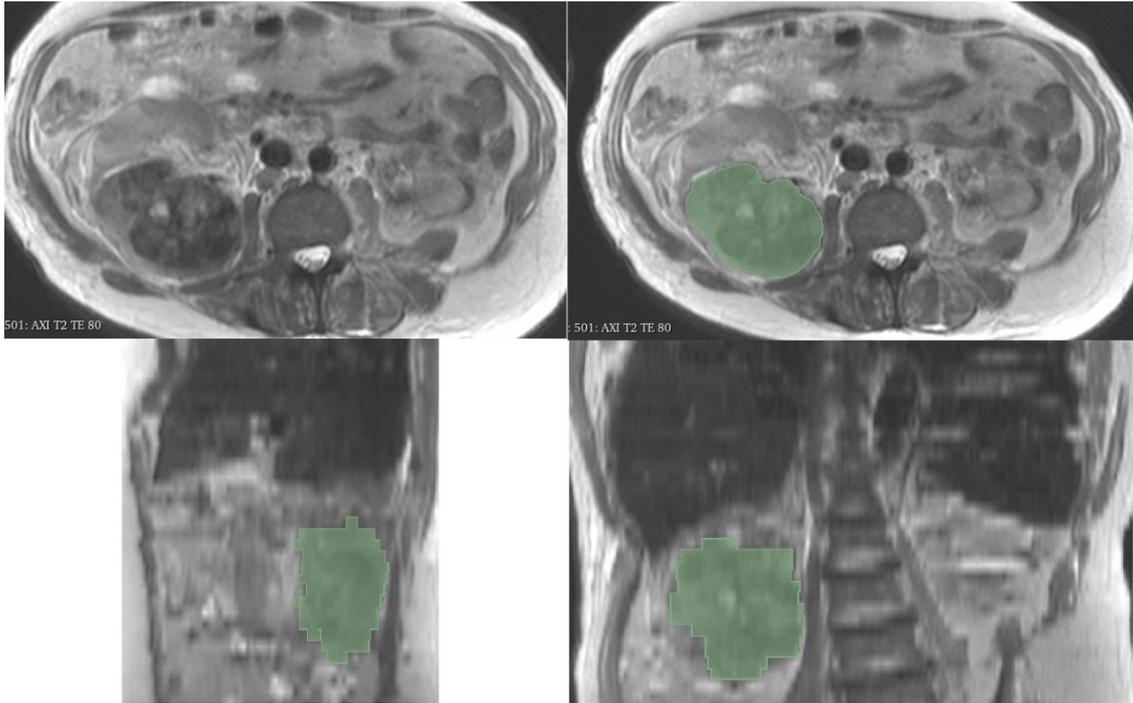
Para a análise radiômica consecutiva, a sequência T2, o mapa ADC (derivado de DWI com valor b de 0/1000) e a sequência T1 pós-contraste na fase portal, foram extraídos no formato “Digital Imaging and Communications in Medicine” (DICOM). A sequência T2 é uma ferramenta útil para avaliação das mais variadas lesões, com excelente análise de caracterização dos diferentes tipos de tecido. Os mapas ADC são uma ferramenta útil para avaliar a difusão das moléculas de água nos tecidos e, portanto, ajudam a identificar áreas de câncer, que geralmente apresentam difusão restrita. As sequências pós-contraste, sobretudo a sequência na fase portal, conseguem avaliar tecidos e lesões que possuem captação após injeção endovenosa.

### **3.8 Segmentação na ressonância magnética**

A segmentação das imagens de RM é uma etapa crucial na análise radiômica, pois permite a identificação e quantificação de características específicas das lesões renais. Em correlação direta com uma estação de trabalho institucional e os respectivos relatórios clínicos, as séries extraídas foram revisadas por um investigador usando a plataforma de software 3D Slicer de código aberto (<http://slicer.org>, versão 4.9.0) [17, 18]. O 3D Slicer é um software avançado e amplamente utilizado na pesquisa médica para visualização e análise de imagens médicas, oferecendo diversas ferramentas para processamento e análise de dados.

Com posicionamento do Volume of Interest (VOI) consecutivo usando imagens das sequências T2, mapa de ADC e T1 pós-contraste, a segmentação manual das lesões foi executada aplicando a ferramenta de pintura do editor de segmentos conforme mostra a Figura 2. A segmentação manual é um processo minucioso, que envolve a demarcação cuidadosa das áreas de interesse em cada fatia sequências avaliadas. Essa etapa é crucial para garantir a precisão na identificação das características radiômicas das lesões renais.

A segmentação foi realizada por radiologista com 1 ano de experiência em radiologia do abdome e pelve, analisando as diferentes sequências e segmentando os nódulos, como demonstrado na Figura 2.



(Fonte: Arquivo próprio)

Figura 2 - Segmentação de lesão renal visualizada na RM no 3D Slicer, com imagens em axial antes e após segmentação, com reconstruções sagital e coronal.

### 3.9 Randomização dos pacientes

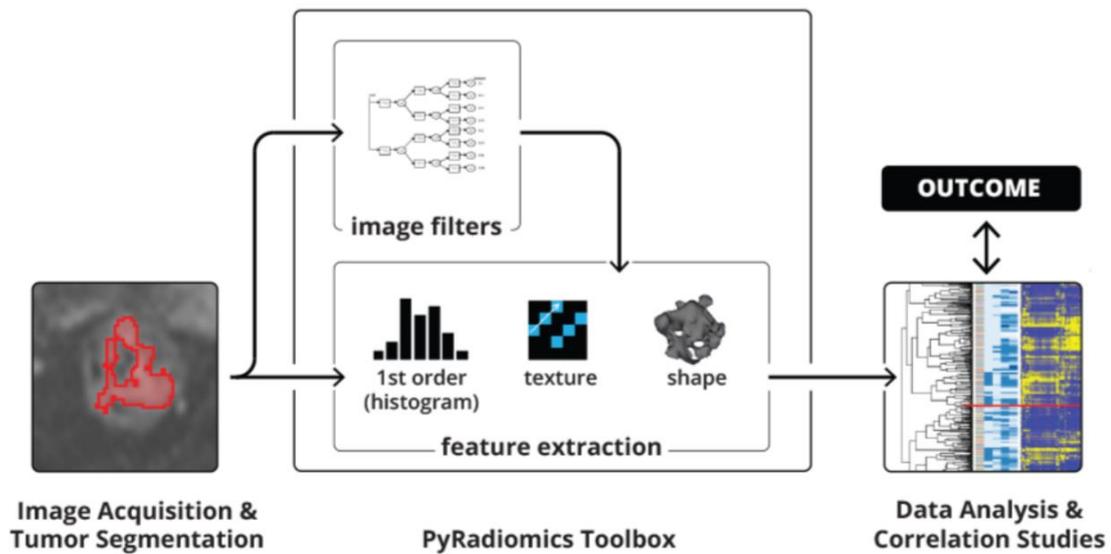
A randomização dos pacientes é um aspecto fundamental no processo de validação do modelo desenvolvido, pois permite a avaliação do desempenho do modelo em um conjunto independente de dados. Isso ajuda a garantir que os resultados obtidos sejam generalizáveis e não sejam apenas aplicáveis ao conjunto de dados utilizado no treinamento do modelo.

Para evitar viés de seleção e garantir a imparcialidade na divisão dos dados, os exames de pacientes foram randomizados em dois grupos: conjunto de treinamento e conjunto de validação. A divisão foi realizada utilizando uma técnica de amostragem aleatória simples, garantindo que cada paciente tivesse a mesma probabilidade de ser incluído em um dos dois grupos. A proporção de pacientes nos conjuntos de treinamento e validação pode variar, mas uma divisão comum é 70% dos dados para treinamento e 30% para validação.

### 3.10 Extração de recursos

Utilizamos o pacote de código aberto PyRadiomics que ganha cada vez mais estabelecimento como padrão de referência para análise radiômica [19]. De sete classes de recursos, todos os recursos padrão foram extraídos: estatísticas de primeira ordem, baseadas em forma, matriz de coocorrência de nível de cinza (GLCM), matriz de comprimento de execução de nível de cinza (GRLLM), matriz de zona de tamanho de nível de cinza (GLSZM), matriz de dependência de nível de cinza (GLDM) e matriz de diferença de tons de cinza vizinhos (NGTDM) levando a 105 características/VOIs em ADC (<http://pyradiomics.readthedocs.io>). Extraímos os recursos com as configurações padrão no PyRadiomics, ou seja, sem reamostragem ou filtragem, sem recursos baseados em wavelets, largura de bin 25 e GLCM simétrico forçado (<http://pyradiomics.readthedocs.io>). A Figura 3 mostra um mapa ilustrativo do funcionamento do PyRadiomics.

A função de processamento de múltiplas pastas foi desenvolvida para iterar sobre todas as pastas contendo imagens médicas e máscaras, chamando a função de extração de características radiômicas para cada par de imagem e máscara. Os resultados das características extraídas foram armazenados em um DataFrame do pandas, com as linhas representando cada exame e as colunas correspondendo às características radiômicas extraídas. Além disso, um identificador único foi atribuído a cada caso, e uma coluna adicional foi incluída no DataFrame para indicar a categoria do tumor (alto ou baixo grau). Após o processamento de todas as pastas e da extração das características radiômicas, o script criou uma pasta para armazenar os resultados em formato CSV. O DataFrame contendo as características e os identificadores foi então exportado como um arquivo CSV, facilitando o uso e a análise dos dados em outras ferramentas e ambientes. Durante o desenvolvimento e a execução do script, foram incluídos comentários e verificações de erros para garantir a correta extração e o correto armazenamento das características radiômicas. Isso incluiu a verificação da existência das pastas e arquivos de imagem e máscara, a validação dos parâmetros de extração e a inspeção visual dos resultados obtidos.



(Fonte: PyRadiomics (<https://www.radiomics.io/pyradiomics.html>))

Figura 3 - Mapa ilustrativo do funcionamento do PyRadiomics.

### 3.11 Análise e processamento dos dados

A análise e o processamento dos dados são etapas cruciais no desenvolvimento e avaliação de modelos de aprendizado de máquina para diagnóstico e prognóstico em imagens médicas. Essas etapas envolvem a manipulação, limpeza, visualização e transformação dos dados para garantir a qualidade, a representatividade e a compreensibilidade das informações obtidas. A seguir, detalhamos as etapas e as decisões tomadas durante a análise e o processamento dos dados neste estudo.

#### 3.11.1 Análise exploratória dos dados

A análise exploratória dos dados permitiu uma compreensão inicial dos padrões e das relações presentes no conjunto de dados. Isso incluiu a geração de estatísticas descritivas, como média, mediana, desvio padrão e distribuição das características, e a visualização dos dados por meio de gráficos, como histogramas, boxplots e scatter plots. Essas análises auxiliaram na identificação de características relevantes, outliers e possíveis correlações entre as variáveis.

### **3.11.2 Redução de dimensionalidade**

A técnica Principal Component Analysis (PCA) foi aplicada para reduzir a dimensionalidade dos dados e identificar as direções de maior variância nos dados. Isso permitiu uma melhor visualização e interpretação dos dados, além de reduzir o tempo de treinamento e melhorar a eficiência dos modelos de classificação. A quantidade de componentes principais a serem mantidos foi selecionada com base na porcentagem acumulada de variância explicada, garantindo a preservação da maior parte das informações presentes nos dados originais.

### **3.11.3 Visualização de dendrogramas e hierarquias de agrupamento**

A visualização de dendrogramas e hierarquias de agrupamento auxiliou na compreensão dos padrões de agrupamento presentes nos dados e na identificação de possíveis subgrupos ou estruturas de dados. Essas visualizações podem ajudar a detectar relações não lineares entre as características e a identificar grupos de pacientes com perfis semelhantes, o que pode ser útil para a personalização do tratamento e a estratificação de risco.

### **3.11.4 Preparação dos dados para treinamento e validação**

O conjunto de dados foi dividido em dois subconjuntos randomicamente: treinamento e validação. Isso é importante para garantir que os modelos de classificação sejam treinados e avaliados em dados representativos e balanceados.

### **3.11.5 Normalização e padronização dos dados**

Os dados foram normalizados e padronizados para garantir que todas as características tivessem a mesma escala e distribuição. Isso é fundamental para garantir a comparabilidade das características e a convergência dos algoritmos de aprendizado de máquina.

Depois de processadas as imagens, os dados de resultado (Features) serão gravados em planilhas de Excel. Os dados nas planilhas (Features) serão processados e consumidos pelos modelos de machine learning. Modelos que serão treinados: *Logistic Regression*, *Random Forest* e *Support Vector Machine*. O modelo com melhor resultado será utilizado.

### **3.11.6 Treinamento e avaliação dos modelos de classificação**

Os modelos de classificação, incluindo Support Vector Machine (SVM), Random Forest e Logistic Regression, foram treinados utilizando o conjunto de dados de treinamento e avaliados com base no conjunto de dados de validação. Métricas de desempenho, como acurácia, matriz de confusão, precisão, recall, F1-score e AUC-ROC, foram calculadas para avaliar o desempenho dos modelos e identificar os mais adequados para a tarefa de classificação.

### **3.11.7 Validação cruzada**

Foi realizada uma validação cruzada com 10 folds para estimar a robustez e a generalização dos modelos de classificação. A validação cruzada envolve a divisão do conjunto de dados em k subconjuntos e a realização do treinamento e validação em cada fold, garantindo uma avaliação mais confiável do desempenho dos modelos. As métricas de desempenho foram calculadas para cada fold e as médias e intervalos de confiança foram reportados.

### **3.11.8 Comparação e seleção do melhor modelo**

Após treinar e avaliar os modelos de classificação e realizar a validação cruzada, os resultados foram comparados para identificar o modelo com o melhor desempenho geral. A seleção do melhor modelo considerou não apenas as métricas de desempenho, mas também a complexidade, a interpretabilidade e a capacidade de generalização dos modelos. Com base nos resultados, o modelo selecionado pode ser usado para classificar novos casos e auxiliar na tomada de decisões clínicas.

### **3.11.9 Implementação e monitoramento do modelo**

Após a validação externa e a seleção do modelo final, o modelo pode ser implementado em sistemas clínicos para auxiliar na classificação de tumores de próstata e na tomada de decisões clínicas. É importante monitorar o desempenho do modelo continuamente e atualizá-lo conforme necessário, incorporando novos dados e refinando as características e hiperparâmetros para garantir a sua eficácia e relevância no ambiente clínico em constante evolução.

## **3.12 Análise estatística:**

Os resultados do estudo foram analisados por meio de estatística descritiva e testes t. Os seguintes passos foram seguidos na análise estatística:

1. A acurácia, precisão e AUC-ROC para cada classificador foram calculados para cada fase, com cálculo de média, desvio-padrão e intervalo de confiança;
2. Em seguida, foi realizado um teste t para comparar o desempenho dos classificadores entre as diferentes fases.

## 4 RESULTADOS

Foram incluídos no estudo 42 pacientes com neoplasia renal, após as exclusões. Entre os 42 pacientes, 4 não fizeram fase portal e 2 não fizeram sequências de difusão/ADC, totalizando respectivamente 38 pacientes no grupo Portal e 40 pacientes no grupo ADC.

Participaram do estudo 29 pacientes (69%) do sexo masculino e 13 pacientes do sexo feminino (31%), com média de idade de 60 anos no sexo masculino (mediana 58 anos, DP 11,9 anos) e 65 anos no sexo feminino (mediana 64 anos, DP 11,1 anos). A distribuição de sexo e idade especificamente para os grupos 1 e 2 está descrita nas Tabelas 1 e 2.

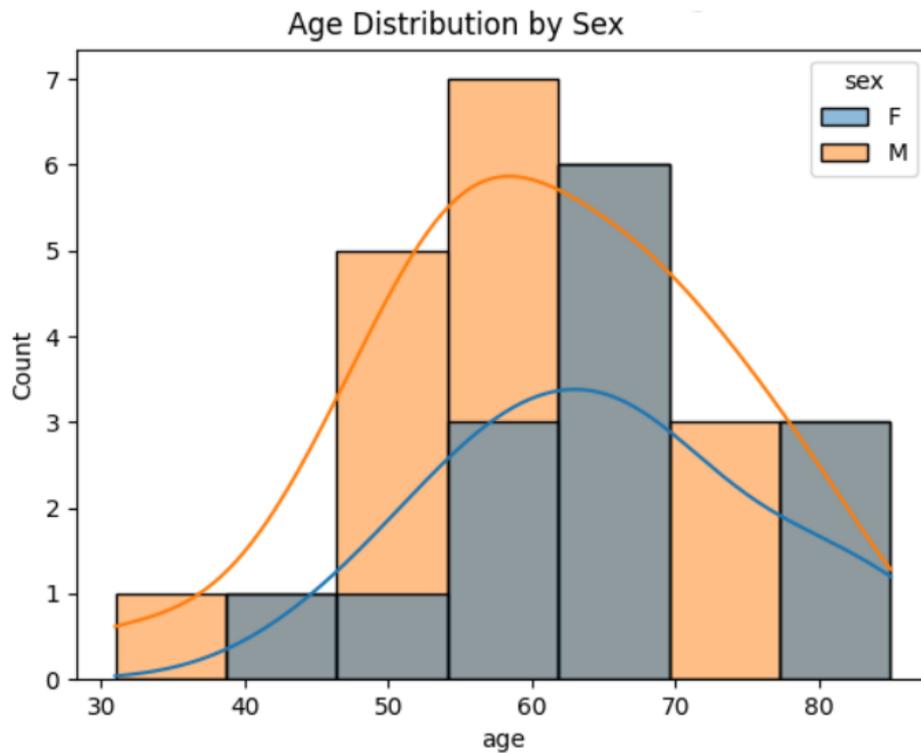
**Tabela 1** - Distribuição dos pacientes por sexo

Grupo	Feminino	Masculino
1 (baixo grau)	10	19
2 (alto grau)	3	10

**Tabela 2** - Distribuição dos pacientes por idade

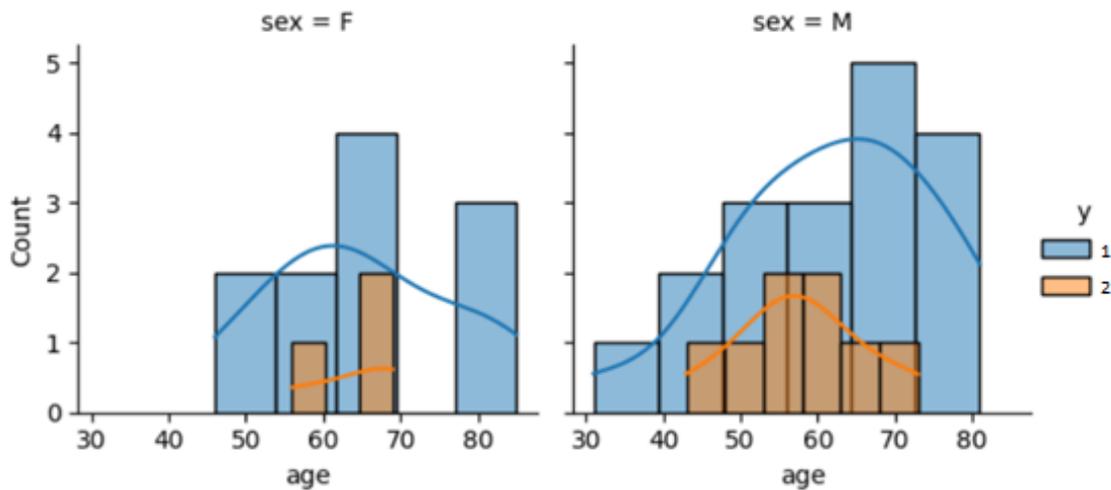
Grupo	Média (anos)	Mediana (anos)	Desvio-padrão (anos)
1 (baixo grau)	62,7	63	12
2 (alto grau)	59,0	57	8,5

A Figura 4 demonstra a distribuição dos tumores por idade e sexo dos pacientes, e a figura 5 demonstra a distribuição dos tumores por idade separada para cada sexo.



Fonte: Arquivo próprio.

Figura 4 - Distribuição dos tumores por idade e sexo.



Fonte: Arquivo próprio. Grupo 1 = baixo grau; Grupo 2= alto grau

Figura 5 - Distribuição por idade separada para cada sexo, por grupo de agressividade anatomopatológica dos tumores.

Dentre os tumores estudados, 8 eram CCR papilíferos, sendo 5 papilíferos tipo 2 e 3 papilíferos tipo 1, outros 2 outros representando CCR papilífero de células claras, além de 2 representando anatomopatológico de CCR inclassificável (ambos classificados como alto grau). Além disso, houveram 2 CCR cromóforos, não incluídos na análise do estudo devido à classificação de Fuhrman não ser recomendada nestes tumores.

Neste estudo, analisamos os dados demográficos de três grupos distintos: ADC, Portal e T2. As estatísticas de avaliação do modelo para cada grupo, incluindo Support Vector Machines (SVM), *Random Forest* e *Logistic Regression*, obtidas usando *cross-validation* de 10-*folds*, são as seguintes:

Para o grupo ADC, o modelo SVM demonstrou uma acurácia média de 0,8000 (95% IC: [0,6441-0,9559]), precisão média de 0,6500 (95% IC: [0,4281-0,8719]) e AUC-ROC média de 0,7333 (95% IC: [0,4553-1,0114]). O desempenho do modelo SVM sugere que ele pode efetivamente distinguir entre tumores renais de baixo e alto grau no conjunto de dados ADC.

Em seguida, avaliamos o modelo *Random Forest* para o conjunto de dados ADC. O modelo alcançou uma acurácia média de 0,6750 (IC de 95%: [0,5140-0,8360]), precisão média de 0,4583 (IC de 95%: [0,2887-0,6280]) e AUC-ROC média de 0,5750 (IC de 95%: [0,3577-0,7923]). O modelo *Random Forest* apresentou desempenho inferior em comparação com o modelo SVM para este conjunto de dados, indicando que pode não ser tão eficaz na classificação dos graus do tumor.

Por fim, avaliamos o modelo de *Logistic Regression* para o grupo ADC. O modelo exibiu uma acurácia média de 0,7750 (IC de 95%: [0,6264-0,9236]), precisão média de 0,5875 (IC de 95%: [0,3757-0,7993]) e AUC-ROC média de 0,7333 (IC de 95%: [0,4553-1,0114]). O modelo de *Logistic Regression* teve desempenho comparável ao modelo SVM, sugerindo que ele também pode fornecer informações valiosas na distinção de graus de tumor no conjunto de dados ADC.

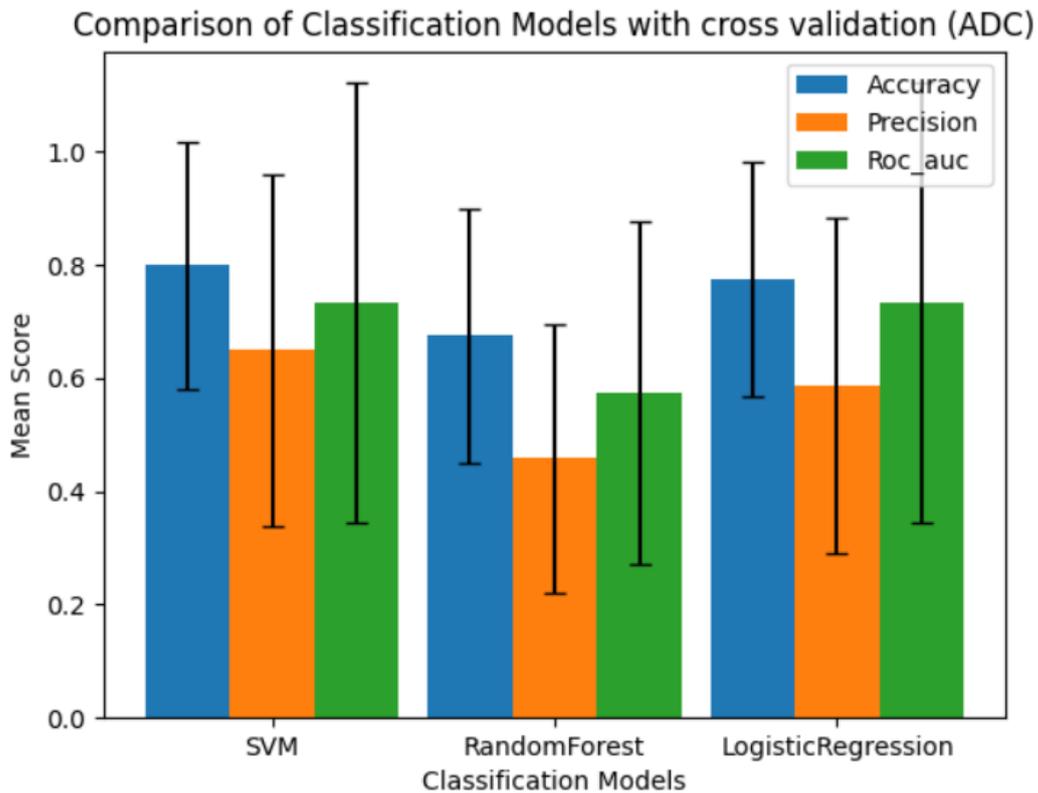
Não houve diferença significativa de acurácia, precisão e AUC-ROC entre os três classificadores para o mapa de ADC.

Os resultados das análises estatísticas para o grupo ADC estão expostos na tabela 3 e a figura 6 demonstra a comparação dos modelos classificadores com *cross validation* de 10 *folds*.

**Tabela 3** - Acurácia, precisão e área sob a curva ROC dos diferentes classificadores utilizados para o grupo ADC

ADC	Acurácia ( $\pm$ DP)	Precisão ( $\pm$ DP)	ROC_AUC ( $\pm$ DP)
SVM	0.80 ( $\pm$ 0.22)	0.65 ( $\pm$ 0.31)	0.73 ( $\pm$ 0.39)
<i>Random Forest</i>	0.68 ( $\pm$ 0.22)	0.46 ( $\pm$ 0.24)	0.73 ( $\pm$ 0.30)
<i>Logistic Regression</i>	0.78 ( $\pm$ 0.21)	0.59 ( $\pm$ 0.30)	0.73 ( $\pm$ 0.39)

SVM: *Support Vector Machine*; DP: desvio-padrão; ADC: *Apparent diffusion coefficient*. AUC: *area under the curve*. ROC: *Receiver operator characteristic*



Fonte:

Arquivo próprio. ADC: *Apparent diffusion coefficient*. SVM: *Support Vector Machine*. AUC: *area under the curve*. ROC: *Receiver operator characteristic*

Figura 6 - Comparação dos modelos classificadores com *cross validation* para o grupo ADC.

Para o grupo Portal, primeiro analisamos o modelo SVM. Atingiu uma acurácia média de 0,8167 (IC de 95%: [0,6980-0,9353]), precisão média de 0,6875 (IC de 95%: [0,4835-0,8915]) e AUC-ROC média de 1,0000 (IC de 95%: [1,0000-1,0000]). O alto desempenho do

modelo SVM indica que ele é altamente eficaz na classificação de graus de tumor dentro do conjunto de dados Portal.

Em seguida, avaliamos o modelo *Random Forest* para o conjunto de dados Portal. O modelo exibiu uma acurácia média de 0,8333 (IC de 95%: [0,7336-0,9331]), precisão média de 0,6167 (IC de 95%: [0,3925-0,8408]) e AUC-ROC média de 0,6667 (IC de 95%: [0,4055-0,9279]). O modelo *Random Forest* exibiu uma precisão ligeiramente melhor do que o modelo SVM, mas teve uma pontuação AUC-ROC mais baixa, sugerindo que pode não ser tão eficaz quanto o SVM na distinção de graus de tumor para o grupo Portal.

Por fim, avaliamos o modelo de *Logistic Regression* para o grupo Portal. O modelo alcançou uma acurácia média de 0,7833 (IC de 95%: [0,6229-0,9437]), precisão média de 0,6417 (IC de 95%: [0,4142-0,8691]) e AUC-ROC média de 0,8167 (IC de 95%: [0,5876-1.0457]). O modelo de *Logistic Regression* teve um desempenho semelhante ao modelo *Random Forest*, com uma precisão ligeiramente inferior, mas uma pontuação AUC-ROC mais alta, sugerindo que ele também pode fornecer informações valiosas para distinguir os graus de tumor no conjunto de dados Portal.

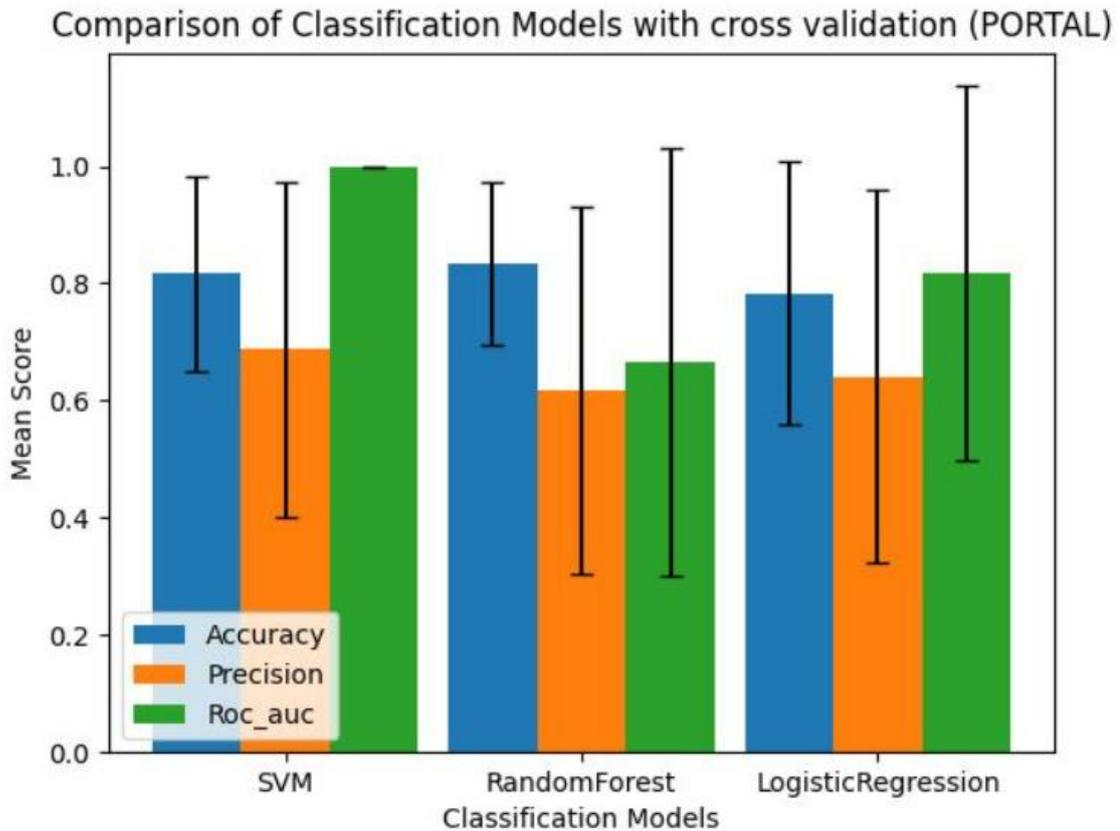
Para o Portal, a AUC-ROC da SVM foi melhor quando comparado com *Logistic Regression*, 1,00 versus 0,82 ( $p=0,002$ ) e quando comparado com *Random Forest*, 1,00 versus 0,67 ( $p<0,00001$ ).

Os resultados das análises estatísticas para o grupo Portal estão expostos na tabela 4 e a figura 7 demonstra a comparação dos modelos classificadores com *cross validation* de 10 *folds*.

**Tabela 4** - Acurácia, precisão e área sob a curva ROC dos diferentes classificadores utilizados para o grupo Portal

Portal	Acurácia ( $\pm$ DP)	Precisão ( $\pm$ DP)	ROC_AUC ( $\pm$ DP)
SVM	0.82 ( $\pm$ 0.17)	0.69 ( $\pm$ 0.29)	1.00 ( $\pm$ 0)
<i>Random Forest</i>	0.83 ( $\pm$ 0.14)	0.62 ( $\pm$ 0.31)	0.67 ( $\pm$ 0.37)
<i>Logistic Regression</i>	0.78 ( $\pm$ 0.22)	0.64 ( $\pm$ 0.32)	0.82 ( $\pm$ 0.32)

SVM: *Support Vector Machine*; DP: desvio-padrão; AUC: *area under the curve*. ROC: *Receiver operator characteristic*



Fonte: Arquivo próprio. SVM: *Support Vector Machine*. AUC: *area under the curve*. ROC: *Receiver operator characteristic*

Figura 7 - Comparação dos modelos classificadores com *cross validation* para o grupo Portal.

Para o grupo T2, examinamos primeiro o modelo SVM. Atingiu uma acurácia média de 0,8600 (IC de 95%: [0,7772-0,9428]), precisão média de 0,7250 (IC de 95%: [0,5180-0,9320]) e AUC-ROC média de 0,8667 (IC de 95%: [0,7498-0,9835]). O alto desempenho do modelo SVM para o conjunto de dados T2 indica que ele é altamente eficaz na classificação de graus de tumor dentro desse grupo.

Em seguida, analisamos o modelo *Random Forest* para o conjunto de dados T2. O modelo demonstrou uma acurácia média de 0,7250 (IC de 95%: [0,6235-0,8265]), precisão média de 0,4717 (IC de 95%: [0,3207-0,6226]) e AUC-ROC média de 0,6333 (IC de 95%: [0,4282-0,8385]). O modelo *Random Forest* mostrou um desempenho inferior em comparação com o modelo SVM para o conjunto de dados T2, indicando que pode não ser tão eficaz na classificação de graus de tumor dentro deste grupo.

Por fim, avaliamos o modelo de *Logistic Regression* para o grupo T2. O modelo exibiu uma precisão média de 0,7850 (IC de 95%: [0,6600-0,9100]), precisão média de 0,7208 (IC de 95%: [0,5390-0,9027]) e AUC-ROC médio de 0,8333 (IC de 95%: [0,7141-0,9526]). O modelo de *Logistic Regression* teve um desempenho semelhante ao modelo SVM, sugerindo

que também pode fornecer informações valiosas na distinção de graus de tumor no conjunto de dados T2.

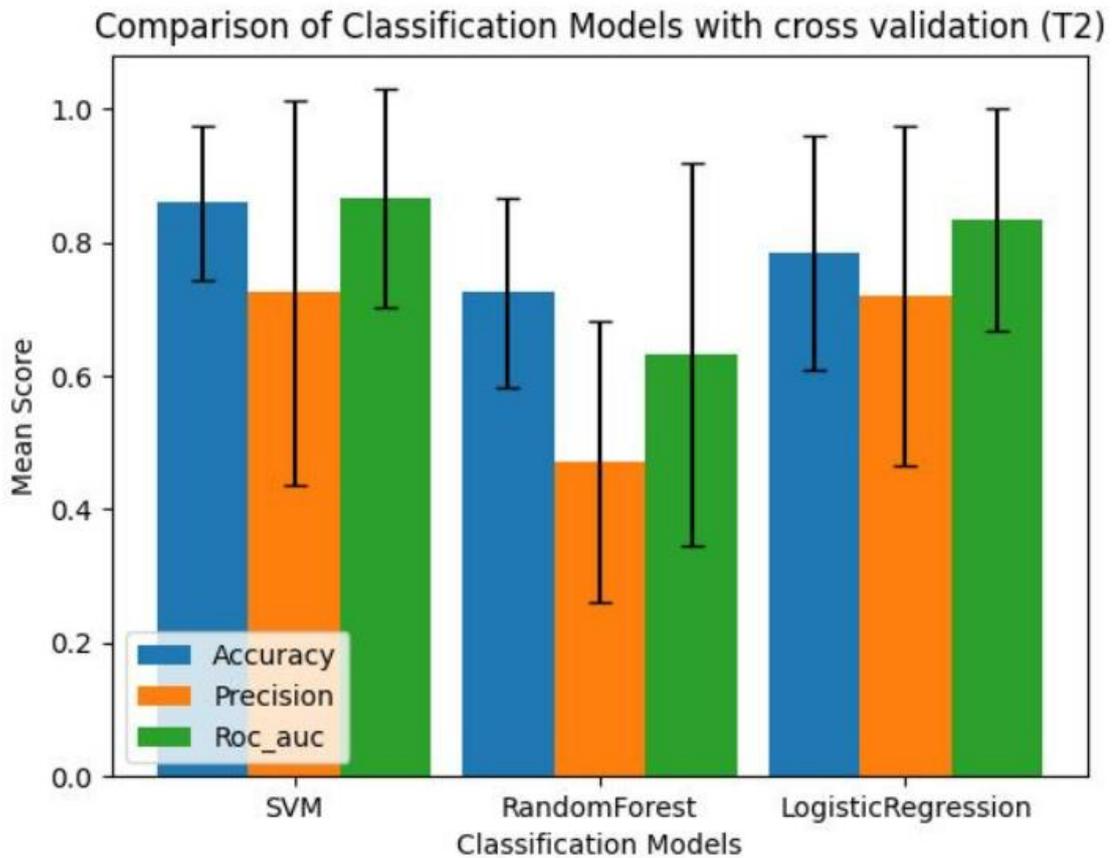
Para o grupo T2, a AUC-ROC da SVM foi melhor quando comparado com *Random Forest*, 1,00 versus 0,67 ( $p < 0,00001$ ), mas não quando comparado com *Logistic Regression*. Para a precisão, houve diferença estatisticamente significativa entre SVM e *Logistic Regression* para *Random Forest* ( $p = 0,01$ ).

Os resultados das análises estatísticas para o grupo T2 estão expostos na tabela 5 e a figura 8 demonstra a comparação dos modelos classificadores com *cross validation* de 10 *folds*.

**Tabela 5** - Acurácia, precisão e área sob a curva ROC dos diferentes classificadores utilizados para o grupo T2

T2	Acurácia ( $\pm$ DP)	Precisão ( $\pm$ DP)	ROC_AUC ( $\pm$ DP)
SVM	0.86 ( $\pm$ 0.12)	0.72 ( $\pm$ 0.29)	0.87 ( $\pm$ 0.16)
<i>Random Forest</i>	0.72 ( $\pm$ 0.14)	0.47 ( $\pm$ 0.21)	0.63 ( $\pm$ 0.29)
<i>Logistic Regression</i>	0.78 ( $\pm$ 0.17)	0.72 ( $\pm$ 0.25)	0.83 ( $\pm$ 0.17)

SVM: *Support Vector Machine*; DP: desvio-padrão; AUC: *area under the curve*. ROC: *Receiver operator characteristic*



Fonte: Arquivo próprio. SVM: *Support Vector Machine*. AUC: *area under the curve*. ROC: *Receiver operator characteristic*

Figura 8 - Comparação dos modelos classificadores com *cross validation* para o grupo T2.

Não houve diferença estatisticamente significativa entre a acurácia de SVM, *Random Forest* e *Logistic Regression* nas diferentes sequências, T2, Portal e ADC. O mesmo aconteceu para a precisão média. No entanto, para a AUC-ROC, o valor obtido para a SVM, na fase portal foi estatisticamente significativo quando comparado àquele do T2, 1,00 versus 0,87 ( $p=0,008$ ) e do ADC, 1,00 versus 0,73 ( $p=0,001$ ).

## 5 DISCUSSÃO

No nosso estudo foram desenvolvidos, validados e comparados modelos de aprendizado de máquina para prever o grau histológico das neoplasias renais utilizando três sequências de RM. Especificamente, avaliamos o desempenho de três modelos de classificação amplamente utilizados: *Support Vector Machines* (SVM), *Random Forest* e *Logistic Regression*. O objetivo era identificar um modelo preditivo eficaz que pudesse ajudar os médicos a tomar decisões mais precisas no manejo de tumores renais.

Os modelos de aprendizado de máquina para avaliação de características radiômicas na imagem por RM são marcadores promissores da agressividade das neoplasias renais nos níveis histopatológicos [20]. Os resultados apresentados neste estudo indicam que é possível utilizar características radiômicas extraídas de imagens médicas para a classificação de tumores de rim de alto e baixo grau com alta acurácia e desempenho consistente. Além disso, os resultados sugerem que a técnica de seleção de características é importante para aprimorar a eficiência dos modelos de classificação. Nossos achados indicam que o SVM na fase portal, foi modelo com melhor desempenho, quando avaliamos a AUC-ROC, superando consistentemente os outros modelos e o SVM das imagens T2 e do mapa de ADC.

Entre os CCR, o subtipo mais comum é o subtipo células claras. Os CCRcc tendem a ser tumores hipervascularizados e podem ter evidência macroscópica de hemorragia e necrose, fato que já foi descrito de forma sagaz por patologistas de eras mais antigas, em 1904 [21]. Microscopicamente eles são definidos pelo citoplasma claro devido ao acúmulo abundante de conteúdo rico em lipídios e glicogênio [22]. Os achados de imagem clássicos refletem estas características, com lesões heterogêneas devido a hemorragia, necrose, cistos e calcificações, com intenso realce pós-contraste. A necrose é mais comum em lesões maiores, geralmente > 4 cm [23]. A taxa de ocorrência e grau de necrose vem sendo descrito como associado a histologia de tumores de alto grau [24, 25].

A variante papilar (CCRp) é o segunda mais comum no adulto, representando até 15% dos casos e possui ainda uma classificação própria com dois subtipos distintos. Na maior parte dos casos, este subtipo tem prognóstico melhor que o CCRcc, notadamente devido ao CCRp tipo 1, que é o subtipo mais comum e geralmente detectado em estágios mais precoces e menores graus de agressividade. Já o CCR tipo 2, em geral, aparece como tumor de alto grau, podendo estar relacionado a metástase linfonodo e, em alguns casos, invasão venosa [23]. Classicamente, CCRp são lesões bem definidas, sólidas e de crescimento lento [26]. São

nódulos mais homogêneos e hipovasculares ao parênquima adjacente, classicamente com sinal baixo nas sequências ponderadas em T2 da RM [24, 25]. Infelizmente, as técnicas atuais de imagem não conseguem diferenciar entre os dois subtipos do CCRp, de forma pré-operatória [27].

O subtipo cromóforo representa 5% dos tumores e carrega o melhor prognóstico entre os CCR, com metástases ocorrendo em até 7% dos casos, apenas [28]. O aspecto de imagem reflete geralmente uma lesão homogênea, hipovascular em relação ao córtex, com realce intermediário entre os subtipos células claras e papilar [26], apresentando-se com hipossinal ou sinal intermediário em T2, podendo ter uma cicatriz central.

Está claro que existem vários subtipos dentro do espectro do CCR, sobretudo após o crescimento dos estudos moleculares detalhando a biologia divergente dos subtipos. Porém, os vários subtipos histológicos podem variar grandemente no seu prognóstico. Por exemplo, inicialmente foi pensado que o CCR papilífero tinha sempre melhor prognóstico que o subtipo células claras [29]. Porém, com a emergência dos dois tipos histológicos da variante papilar, a história mudou de direção, com CCRp tipo 2 apresentando altas taxas de metástases e pior sobrevida [29], inclusive similares ao CCRcc de alto grau. Neste sentido, como dito acima, as técnicas atuais não conseguem diferenciar com precisão os tipos 1 e 2 da variante papilífera. Kuthi *et al* (2016), num estudo com 928 pacientes, encontrou sobrevida de 5 anos de 100% para os cromóforos, 96% para os CCRcc de baixo grau, 92% para os CCRp tipo 1, 63% para os CCRcc de alto grau e 65% para os CCRp tipo 2 [30]. Deste modo, mesmo a classificação histológica é insuficiente para prever o comportamento do tumor, pois há CCRcc com comportamento indolente e há CCRp (e outros subtipos) com comportamento agressivo [31].

Os resultados deste estudo demonstram que a radiômica pode ser usada para diferenciar entre tumores renais de baixo e alto grau usando dados de RM com boa precisão. A SVM, *Random Forest* e *Logistic Regression* obtiveram bons resultados, com a SVM alcançando a maior AUC-ROC na sequência Portal. Esses achados são consistentes com estudos anteriores que mostraram que características radiômicas extraídas de dados de RM podem ser usadas para diferenciar entre CCR de baixo e alto grau.

Goyal *et al* (2019), num estudo com 34 tumores de CCR encontraram resultados similares na diferenciação de tumores de baixo e alto grau, com melhores resultados utilizando o DWI b1000 e fases corticomedular e nefrográfica pós-contraste. Também demonstraram a capacidade de diferenciar CCR de células claras dos demais subtipos não-células claras. Todas as análises do estudo de Goyal *et al* obtiveram  $AUC > 0.8$  [32]. Zhao *et al* (2020) numa coorte multicêntrica num período de 10 anos estudaram 430 lesões com RM e

encontraram acurácia de 88% no *test set* utilizando a graduação Fuhrman e 83% com o *test set* com ISUP/WHO [33]. Cui *et al* (2020) também encontraram desempenho satisfatório, reprodutível e confiável analisando modelos de radiômica baseados em RM e TC em diferenciar tumores de alto e baixo grau, e nenhum dos modelos superou esmagadoramente o outro, com acurácia na validação interna de 73% para RM e 79% para TC e acurácia na validação externa de 74% na RM e 69% na TC [34]. Há ainda vários outros estudos que demonstraram bons resultados utilizando TC para avaliação de diferenciação de tumores de baixo e alto grau [16, 35, 36], inclusive analisando características da gordura adjacente ao tumor nessa diferenciação, com bons resultados [37].

Apesar da TC ser historicamente o pilar do diagnóstico por imagem do CCR, a RM se tornou uma grande ferramenta para resolução de problemas para lesões renais. Devido a seu melhor contraste para tecidos de partes moles, a RM superou a TC por exemplo na avaliação de massas císticas indeterminadas (Bosniak IIF e III, malignidade em 10% e 50%, respectivamente) [38], invasão local e extensão intravascular [39]. A RM apresenta múltiplos diferentes parâmetros para avaliação tanto da morfologia (sequências T2 e T1 pós-contraste) quanto no microambiente funcional (DWI, perfusão), sendo, portanto, mais provável que possa prover métricas mais robustas e viáveis na análise textural em comparação com a TC [32]. Porém no âmbito da radiômica e suas possíveis padronizações, a RM apresenta uma importante desvantagem em comparação com a TC: a sua produtividade é menor devido a dificuldade de padronização de protocolos de aquisição, levando a heterogeneidade dos dados de imagem [32, 40]. Ainda, a natureza não quantitativa das sequências ponderadas em T1 e T2 da RM introduzem heterogeneidade adicional mesmo quando os parâmetros de aquisição são constantes [40].

Nossos achados indicam que o modelo SVM superou consistentemente os outros modelos, na AUC-ROC, em todas as sequencias, T2, ADC e Portal, sendo o AUC-ROC do SVM na fase portal, superior aos congêneres ADC e T2. Isto sugere robustez e eficácia no manejo dos recursos radiômicos derivados de RM abdominal. Os resultados também destacam o potencial da *Logistic Regression* na classificação de tumores renais. Em contraste, *Random Forest* demonstrou desempenho relativamente inferior em todos os conjuntos de dados. Isso pode ser devido à presença de recursos ruidosos, à sensibilidade do modelo ao ruído ou ao *overfitting*.

Vale a pena notar que os desvios padrão para todos os três classificadores foram relativamente altos, indicando que pode haver alguma variabilidade no desempenho dos classificadores. Isso pode ser devido ao pequeno tamanho da amostra do estudo, pois estudos

anteriores mostraram que tamanhos de amostra maiores estão associados a maior exatidão e precisão em estudos radiômicos [34]. Além disso, a variabilidade no desempenho dos classificadores também pode ser devida à heterogeneidade inerente dos CCRs, pois os tumores podem variar em suas características microestruturais mesmo dentro do mesmo grau [23]. Uma investigação mais aprofundada com dados adicionais, técnicas de engenharia de recursos ou algoritmos alternativos de aprendizado de máquina pode render resultados ainda melhores. Além disso, a integração do conhecimento específico do domínio, como informações clínicas ou patológicas, pode melhorar ainda mais as capacidades preditivas dos modelos.

Em meio a natureza heterogênea dos CCR, é importante ressaltar que, apesar do ímpeto dos estudos em fazer análises com inteligência artificial na diferenciação histológica entre os tipos de CCR e entre os tumores de baixo e alto grau, um dos parâmetros mais importantes a serem avaliados é o impacto que tais análises podem ter na prática clínica futuramente, ou seja, o quanto essas ferramentas serão capazes de prever, de fato, desfechos clínicos desfavoráveis, e quais seriam as características destas análises com maior capacidade de predição destes desfechos. Um dos desdobramentos dos estudos vigentes do nosso grupo será tentar avaliar, através de análises com inteligência artificial, quais informações podem definir desfechos clínicos desfavoráveis, dado que não foi avaliado neste estudo, tampouco nos estudos publicados até a presente data.

Este estudo possui algumas limitações. Uma limitação potencial deste estudo é que os dados de RM foram adquiridos usando um scanner de 1,5 Tesla, o que pode não ser representativo dos dados de RM adquiridos usando outros scanners. Sabe-se que a intensidade do campo dos scanners de RM pode afetar a precisão dos recursos radiômicos [41]. Portanto, seria interessante replicar este estudo usando dados de RM adquiridos em diferentes scanners para determinar a generalização dos achados. Outra limitação é a amostra do estudo, com número final pequeno de 42 tumores e com distribuição distorcida dos subgrupos. Ainda, sabemos que, embora a RM forneça maiores parâmetros para análise tecidual em comparação com a TC, a reprodutibilidade da análise textural com RM é menor devido à falta de protocolos de aquisição padronizados.

## 6 CONCLUSÃO

Este estudo teve como objetivo avaliar a capacidade da radiômica em diferenciar entre tumores renais de baixo e alto grau usando dados de RM de 42 pacientes. Os resultados do estudo mostraram que a radiômica pode ser usada para diferenciar entre tumores renais de baixo e alto grau com boa precisão usando dados de RM. O SVM e *Logistic Regression* obtiveram bons resultados, com o SVM alcançando os melhores resultados na sequência Portal, ao passo que o *Random Forest* não obteve bons resultados. Esses achados são consistentes com estudos anteriores que mostraram que características radiômicas extraídas de dados de ressonância magnética podem ser usadas para diferenciar entre CCRs de baixo e alto grau.

Em conclusão, nosso estudo contribui para o crescente corpo de pesquisa sobre a aplicação de aprendizado de máquina e radiômica no campo da oncologia. Os resultados promissores obtidos dos modelos SVM e *Logistic Regression* oferecem informações valiosas para pesquisadores e profissionais que buscam desenvolver modelos preditivos eficazes para a classificação de tumores renais usando recursos radiômicos de imagens de ressonância magnética. Estudos futuros podem explorar o uso de outros algoritmos de classificação, técnicas de aprendizado profundo ou a incorporação de dados de imagem multimodais para aprimorar ainda mais o desempenho e a aplicabilidade dos modelos em ambientes clínicos.

## 7 REFERÊNCIAS BIBLIOGRÁFICAS

1. F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *Ca-Cancer J. Clin.* 68 (2018) 394-424.
2. J.C. Cheville, C.M. Lohse, H. Zincke, A.L. Weaver, M.L. Blute, Comparisons of outcome and prognostic features among histologic subtypes of renal cell carcinoma, *Am. J. Surg. Pathol.* 27 (2003) 612-624
3. W. Street, Cancer facts & figures 2019, *Am. Canc. Soc.* 76 (2018)
4. H. Moch, A.L. Cubilla, P.A. Humphrey, V.E. Reuter, T.M. Ulbright, The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours, *European urology* 70(1) (2016) 93-105
5. B.R. Lane, M.K. Samplaski, B.R. Herts, M. Zhou, A.C. Novick, S.C. Campbell, Renal mass biopsy - a renaissance?, *The Journal of urology* 179(1) (2008) 20-7.
6. Delahunt B, Srigley JR, Montironi R, *et al.* Advances in renal neoplasia: recommendations from the 2012 International Society of Urological Pathology Consensus Conference. *Urology.* 2014;83:969-974.
7. GILLIES, R. J.; KINAHAN, P. E.; HRICAK, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*, v. 278, n. 2, p. 563-577, fev. 2016.

8. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018 Aug;18(8):500-510.
9. Shiri I, Maleki H, Hajianfar G, Abdollahi H, Ashrafinia S, Hatt M, *et al*. Next Generation Radiogenomics Sequencing for Prediction of EGFR and KRAS Mutation Status in NSCLC Patients Using Multimodal Imaging and Machine Learning Approaches. *arXiv preprint arXiv:190702121*. 2019
10. EUROPEAN SOCIETY OF RADIOLOGY (ESR). What the radiologist should know about artificial intelligence - an ESR white paper. *Insights into Imaging*, v. 10, n. 1, 4 abr. 2019.
11. Castellano G., Bonilha L., Li L.M., Cendes F. Texture analysis of medical images. *Clin. Radiol*. 2004;59:1061-1069.
12. Mühlbauer J, Egen L, Kowalewski KF, Grilli M, Walach MT, Westhoff N, Nuhn P, Laqua FC, Baessler B, Kriegmair MC. Radiomics in Renal Cell Carcinoma-A Systematic Review and Meta-Analysis. *Cancers (Basel)*. 2021 Mar 17;13(6):1348.
13. Bauman T.M., Potretzke A.M., Wright A.J., Knight B.A., Vetter J.M., Figenshau R.S. Partial Nephrectomy for Presumed Renal-Cell Carcinoma: Incidence, Predictors, and Perioperative Outcomes of Benign Lesions. *J. Endourol*. 2017;31:412-417.
14. T. Hodgdon, M.D. McInnes, N. Schieda, T.A. Flood, L. Lamb, R.E. Thornhill, Can Quantitative CT Texture Analysis be Used to Differentiate Fat-poor Renal

- Angiomyolipoma from Renal Cell Carcinoma on Unenhanced CT Images?, *Radiology* 276(3) (2015) 787-96.
15. Q. Yin, S.C. Hung, L. Wang, W. Lin, J.R. Fielding, W.K. Rathmell, A.H. Khandani, M.E. Woods, M.I. Milowsky, S.A. Brooks, E.M. Wallen, D. Shen, Associations between Tumor Vascularity, Vascular Endothelial Growth Factor Expression and PET/MRI Radiomic Signatures in Primary Clear-Cell-RenalCell-Carcinoma: Proof-of-Concept Study, *Scientific reports* 7 (2017) 43356.
  16. Feng Z, Shen Q, Li Y, Hu Z. CT texture analysis: a potential tool for predicting the Fuhrman grade of clear-cell renal carcinoma. *Cancer Imaging*. 2019;19(1):6.
  17. KUMAR, V. *et al.* Radiomics: the process and the challenges. *Magnetic Resonance Imaging*, v. 30, n. 9, p. 1234-1248, nov. 2012.
  18. FEDOROV, A. *et al.* 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging*, v. 30, n. 9, p. 1323-1341, nov. 2012.
  19. BLEKER, J. *et al.* Multiparametric MRI and auto-fixed volume of interest-based radiomics signature for clinically significant peripheral zone prostate cancer. *European Radiology*, v. 30, n. 3, p. 1313-1324, 27 nov. 2019.
  20. Mühlbauer J, Egen L, Kowalewski KF, Grilli M, Walach MT, Westhoff N, Nuhn P, Laqua FC, Baessler B, Kriegmair MC. Radiomics in Renal Cell Carcinoma-A Systematic Review and Meta-Analysis. *Cancers (Basel)*. 2021 Mar 17;13(6):1348.

21. Jellinek EO. Grawitz Tumor of Kidney. *Cal State J Med.* 1904;2:78-79.
22. Decastro GJ, McKiernan JM. Epidemiology, clinical staging, and presentation of renal cell carcinoma. *Urol Clin North Am.* 2008;35:581-592.
23. Muglia VF, Prando A. Renal cell carcinoma: histological classification and correlation with imaging findings. *Radiol Bras.* 2015 May-Jun;48(3):166-74.
24. Pedrosa I, Sun MR, Spencer M, *et al.* MR imaging of renal masses: correlation with findings at surgery and pathologic analysis. *Radiographics.* 2008;28:985-1003.
25. Prasad SR, Humphrey PA, Catena JR, *et al.* Common and uncommon histologic subtypes of renal cell carcinoma: imaging spectrum with pathologic correlation. *Radiographics.* 2006;26:1795-1806.
26. Lubensky IA, Schmidt L, Zhuang Z, *et al.* Hereditary and sporadic papillary renal carcinomas with c-met mutations share a distinct morphological phenotype. *Am J Pathol.* 1999;155:517-526.
27. Yoshimitsu K, Kakihara D, Irie H, *et al.* Papillary renal carcinoma: diagnostic approach by chemical shift gradient-echo and echo-planar MR imaging. *J Magn Reson Imaging.* 2006;23:339-344.
28. Lopez-Beltran A, Scarpelli M, Montironi R, *et al.* 2004 WHO classification of renal tumours of the adults. *Eur Urol.* 2006;49:798-805.

29. Mancilla-Jimenez R, Stanley RJ, Blath RA. Papillary renal cell carcinoma: a clinical, radiologic, and pathologic study of 34 cases. *Cancer*. 1976;38:2469-2480.
30. Linehan WM, Spellman PT, Ricketts CJ, *et al*. Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma. *N Engl J Med*. 2016;374:135-145.
31. Kuthi L, Jenei A, Hajdu A, Németh I, Varga Z, Bajory Z, Pajor L, Iványi B. Prognostic Factors for Renal Cell Carcinoma Subtypes Diagnosed According to the 2016 WHO Renal Tumor Classification: a Study Involving 928 Patients. *Pathol Oncol Res*. 2017 Jul;23(3):689-698.
32. Goyal A, Razik A, Kandasamy D, Seth A, Das P, Ganeshan B, Sharma R. Role of MR texture analysis in histological subtyping and grading of renal cell carcinoma: a preliminary study. *Abdom Radiol (NY)*. 2019 Oct;44(10):3336-3349.
33. Zhao Y, Chang M, Wang R, *et al*. Deep learning based on MRI for differentiation of low and high-grade in low-stage renal cell carcinoma. *J Magn Reson Imaging* 2020;52:1542-9.
34. Cui, E., Li, Z., Ma, C. *et al*. Predicting the ISUP grade of clear cell renal cell carcinoma with multiparametric MR and multiphase CT radiomics. *Eur Radiol* 30, 2912-2921 (2020).
35. He X, Wei Y, Zhang H, Zhang T, Yuan F, Huang Z, Han F, Song B. Grading of Clear Cell Renal Cell Carcinomas by Using Machine Learning Based on Artificial Neural Networks and Radiomic Signatures Extracted From Multidetector Computed Tomography Images. *Acad Radiol*. 2020 Feb;27(2):157-168.

36. Lin, F., Cui, EM., Lei, Y. *et al.* CT-based machine learning model to predict the Fuhrman nuclear grade of clear cell renal cell carcinoma. *Abdom Radiol* 44, 2528-2534 (2019).
37. Gill TS, Varghese BA, Hwang DH, Cen SY, Aron M, Aron M, Duddalwar VA. Juxtatumoral perinephric fat analysis in clear cell renal cell carcinoma. *Abdom Radiol (NY)*. 2019 Apr;44(4):1470-1480.
38. Defortescu G, Cornu J-N, Béjar S *et al* (2017) Diagnostic performance of contrast-enhanced ultrasonography and magnetic resonance imaging for the assessment of complex renal cysts: a prospective study. *Int J Urol* 24:184-189
39. Karlo CA, Di Paolo PL, Donati OF *et al* (2013) Renal cell carcinoma: role of MR imaging in the assessment of muscular venous branch invasion. *Radiology* 267:454-459
40. Ursprung S, Beer L, Bruining A, Woitek R, Stewart GD, Gallagher FA, Sala E. Radiomics of computed tomography and magnetic resonance imaging in renal cell carcinoma-a systematic review and meta-analysis. *Eur Radiol*. 2020 Jun;30(6):3558-3566.
41. Ammari S, Pitre-Champagnat S, Dercle L, Chouzenoux E, Moalla S, Reuze S, Talbot H, Mokoyoko T, Hadchiti J, Diffetocq S, Volk A, El Haik M, Lakiss S, Balleyguier C, Lassau N, Bidault F. Influence of Magnetic Field Strength on Magnetic Resonance Imaging Radiomics Features in Brain Imaging, an In Vitro and In Vivo Study. *Front Oncol*. 2021 Jan 20;10:541663.