UNIVERSIDADE DE SÃO PAULO
FFCLRP – DEPARTAMENTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA APLICADA À
MEDICINA E BIOLOGIA

**“Análise de textura em pacientes com lesão de glioblastoma
multiforme para correlação com o prognóstico da doença”**

Pedro Henrique de Marco Borges

Dissertação apresentada à Faculdade de Filosofia,
Ciência e Letras de Ribeirão Preto da USP, como parte
das exigências para a obtenção do título de Mestre em
Ciências, Área: Física Aplicada à Medicina e Biologia.

Ribeirão Preto – SP

2020

Pedro Henrique de Marco Borges

**Análise de textura em pacientes com lesão de glioblastoma multiforme
para correlação com o prognóstico da doença**

Versão corrigida

Dissertação apresentada à Faculdade de Filosofia,
Ciência e Letras de Ribeirão Preto da USP, como parte
das exigências para a obtenção do título de Mestre em
Ciências,

Área: Física Aplicada à Medicina e Biologia.

Orientadora: Prof^ª Dr^ª Juliana Fernandes Pavoni

Ribeirão Preto – SP

2020

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.

Borges, Pedro Henrique de Marco

Análise de textura em paciente com lesão de glioblastoma multiforme para correlação com o prognóstico da doença. Ribeirão Preto, 2020.

(90 páginas) p. : il. ; 30 cm

Dissertação de Mestrado, apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto/USP. Área de concentração: Física Aplicada à Medicina e Biologia.

Orientadora: Pavoni, Juliana Fernandes.

1. GBM. 2. Endpoints 3. Teste-t 4. Características 5. Radiomics.

Nome: Pedro Henrique de Marco Borges

Título: Análise de textura em pacientes com lesão de glioblastoma multiforme para correlação com o prognóstico da doença

Dissertação apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo, como requisito para obtenção do título de Mestre em Ciências.

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____

Avaliação: _____

Prof. Dr. _____

Instituição: _____

Avaliação: _____

Prof. Dr. _____

Instituição: _____

Avaliação: _____

In memoriam: Dedé e Sanfoneiro

AGRADECIMENTOS

Camilla

Cláudia

Gustavo Viani

Jéssica

João Henrique

Juliana Pavoni

Léo

Romilda

Zeco

Recursos

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

RESUMO

BORGES, Pedro Henrique de Marco. **Análise de textura em paciente com lesão de glioblastoma multiforme para correlação com o prognóstico da doença.** Ribeirão Preto, 2020. Dissertação (Mestrado em Física Aplicada à Medicina e Biologia) – Departamento de Física, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2020.

Tumores sólidos são temporalmente heterogêneos. O padrão desta heterogeneidade tem sido estudado através da extração de características quantitativas de imagens médicas sofisticadas, como as imagens de ressonância magnética (IRM). Essas características podem se relacionar com os padrões genéticos do tumor e consequentemente com a resposta ao tratamento. Um dos exemplos mais estudados de tumor com alta heterogeneidade é o glioblastoma multiforme (GBM). Essa característica limita o uso de biópsias invasivas para caracterização do tecido, porém, diversos estudos sobre a caracterização da heterogeneidade dessas lesões de forma não invasiva têm sido realizados. GBM possui um mau prognóstico devido à alta recorrência da lesão. A sobrevida média dos pacientes é de 14 meses, quando o tratamento padrão é realizado. Diante deste cenário, é necessário que se faça uma precisa escolha do tratamento para não prejudicar a qualidade de vida do paciente. Neste contexto, o presente trabalho tem como objetivo principal extrair informações quantitativas de IRM de pacientes com GBM que foram submetidos ao tratamento radioterápico, através de processamento de imagens para correlaciona-las com determinados prognósticos da doença. A metodologia desse estudo consistiu em extrair 105 características quantitativas por Radiomics do volume de tumor presente em IRM ponderadas em T1 após administração de contraste de 43 pacientes submetidos à radioterapia no Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto. O software usado para essa extração foi o 3D Slicer – um software gratuito de processamento de imagens. Além disso, os desfechos clínicos ou *endpoints* desses pacientes em termos de tempo para sobrevida global (SG), sobrevida livre da progressão da doença (SLPD) e recorrência na borda do campo de tratamento (RECURR) também foram coletados. Testes estatísticos (Shapiro-Wilk e teste t de *Student*) foram aplicados para determinar quais características se relacionavam de maneira significativa com determinados *endpoints*. Como resultado observou-se que, para o endpoint SLPD em até 3 meses, a característica Radiomics Kurtosis obteve o melhor resultado na separação dos grupos de pacientes com desfecho clínico 1 e 0 (representação binária para pacientes que recidivaram e não recidivaram dentro de 3 meses, respectivamente), com uma acurácia de 69%. Para essa mesma característica, aplicou-se um classificador baseado em Regressão Logística e obteve-se uma área sobre a curva - AUC - de 75%. Com base na revisão bibliográfica de trabalhos envolvendo *endpoints*, Radiomics e GBM, replicamos um trabalho no qual utilizou uma técnica de redução de dimensionalidade baseada na correlação de Spearman para criação de um modelo preditivo utilizando árvores de decisão e obtivemos um resultado consistente com o proposto de 62,8% de acurácia no modelo preditivo para o *endpoint* RECURR.

Palavras chave: Glioblastoma Multiforme. Radiomics. Kurtosis. Aprendizado de máquina.

ABSTRACT

BORGES, Pedro Henrique de Marco. **Texture analysis in patients with glioblastoma multiform lesion to correlation with disease prognosis.** Ribeirão Preto, 2020. Dissertation (Master's degree in physics Applied in Medicine and Biology) – Department of Physics, Faculty of Philosophy, Sciences and Letters of Ribeirão Preto, University of São Paulo, 2020.

Solid tumors are spatial and temporally heterogeneous. Its pattern has been studied through quantitative features extraction from medical images, as magnetic resonance imaging (MRI). These features can be related with genetic patterns of tumors and hence with treatment response. The most studied example with high heterogeneity is glioblastoma multiform (GBM). This characteristic limit the use of invasive biopsies to tissue characterization, but few studies about GBM characterization with a non-invasive way has been done. GBM has a bad prognosis due to high tumor recurrence. The average survival is around 14 months, when the default treatment is selected. That said, a very precise choice for treatment must be done to avoid decreasing patient's performance. In this context, this study has the main goal extract quantitative features through MRI from GBM patients that were submitted to radiotherapy with image processing to correlate them with specific GBM prognosis. The methodology consisted in 105 features extracted by Radiomics analysis from a region of interest (ROI) in T1-post contrast-weighted-MRI of 43 patients submitted to radiotherapy at Clinics Hospital of Faculty of Medicine of Ribeirão Preto. The software used in feature extraction was 3D Slicer – an open source for image processing. Beside this, the endpoints Overall Survival (SG), Progression Free Survival (SLPD) and Lesion Recurrence (RECURR) inside treatment field were collected. Statistical tests (Shapiro-Wilk and t-Student) were applied to achieve which features were significantly correlated with endpoints. As result, for 3 months SLPD, the Kurtosis feature has shown the best result in separating patients with clinical endpoints 1 and 0 (binary representation for patients who had lesion recurrence within 3 months or not, respectively), with an accuracy of 69%. For this same variable, a Logistic Regression classifier was applied and an area under curve - AUC - of 75% was observed in this model. With respect on literature, studies involving Radiomics, GBM and clinical endpoints, we reproduced a methodology in which used a dimensionality reduction technique based on Spearman's rank to create a predictive model using Decision Trees and we got a similar result when comparing the accuracy of 62,8% for RECURR endpoint.

Keywords: Glioblastoma Multiform. Radiomics. Kurtosis. Machine Learning.

LISTA DE ILUSTRAÇÕES

Figura 2.1: Representação do GBM (tumor ativo e edemas).....	24
Figura 2.2: Aplicação simplificada de modelos preditivos para correlação com o prognóstico do padrão celular.....	26
Figura 3.1: Representação do fluxo de trabalho da técnica Radiomics.....	29
Figura 4.1: Representação gráfica do modelo de redução de peso, com a regressão linear.....	32
Figura 4.2: Exemplos de classificação de tipos de pimenta.....	35
Figura 4.3: Fluxograma de aprendizado de máquina.....	36
Figura 4.4: Representação abstrata de uma árvore de decisões.....	36
Figura 4.5: Representação mais precisa de uma DT.....	37
Figura 4.6: Exemplo de um trecho de fluxograma.....	38
Figura 4.7: Exemplo de decisão baseada em características: gato ou centopeia.....	39
Figura 4.8: Resultados de classificação usando DT em forma de clusters.	40
Figura 4.9: Apresentação gráfica dos resultados da classificação usando DT	41
Figura 4.10: Representação gráfica da função sigmoide.....	43
Figura 4.11: Resultado do k-means, com 3 clusters ou agrupamentos.....	45
Figura 4.12: Aplicação do Elbow Method no algoritmo K-means.....	46
Figura 5.1: Representação do delineamento das estruturas GTV e CTV em uma IRM. (a) IRM sem conjunto de estruturas; (b) IRM com conjunto de estruturas	48
Figura 5.2: Exemplo de avaliação da D95% no DVH para GTV/CTV.....	52
Figura 6.1: Gráfico de dispersão da característica SLPD com faixa temporal Mediana 58	
Figura 6.2: Gráfico de dispersão para a característica SLPD com a faixa temporal de 3 meses e o limiar da Kurtosis em 0,4 (escala logarítmica.....	59

Figura 6.3: Matriz de confusão para o endpoint SLPD com a faixa temporal de 3 meses, feita a partir da análise estatística para estimar a acurácia dessa separação.....	60
Figura 6.4: : Curva ROC e matriz de confusão para o modelo preditivo baseado em LR para SLPD em 3 meses.....	61
Figura 6.5: Gráfico de dispersão para endpoint SLPD com faixa temporal de 6 meses. 62	
Figura 6.6: Gráfico de dispersão para endpoint SLPD com faixa temporal de 12 meses....	64
Figura 6.7: Gráfico de dispersão para o endpoint SLPD com faixa temporal de 18 meses. 66	
Figura 6.8: Gráfico de dispersão para o endpoint SLPD com faixa temporal de 24 meses 68	
Figura 6.9: Gráfico de dispersão para o endpoint SG com faixa temporal de 12 meses. O destaque em vermelho representa um paciente classificado como outlier.....	70
Figura 6.10: Gráfico de dispersão para o endpoint SG com faixa temporal de 30 meses...	72
Figura 6.11: Gráfico de dispersão de pares para o endpoint RECURR.....	75
Figura 6.12: Curva ROC para o aprendizado de máquina utilizando a Regressão Logística como algoritmo preditor.....	77
Figura 6.13: Matriz de Confusão para o aprendizado de máquina utilizando a Regressão Logística como algoritmo preditor.....	77
Figura 6.14: Curva ROC para a segunda abordagem utilizando o ranking de Spearman para o endpoint RECURR.....	78
Figura 6.15: Curva ROC para a segunda abordagem utilizando o endpoint SLPD com a faixa temporal mediana.....	79
Figura 6.16: Curva ROC para a segunda abordagem utilizando o endpoint SLPD com a faixa temporal mediana.....	80

LISTA DE TABELAS

Tabela 2.1: Definição de gliomas de acordo com a Organização Mundial da Saúde....	23
Tabela 3.1: Resumo de metodologia e resultados para diferentes abordagens de aprendizado de máquina em pacientes com GBM.....	28
Tabela 4.1: Dados para exemplificar a construção de um modelo a partir de experimentos e porcentagem de redução de peso.....	32
Tabela 5.1: Representação dos endpoints.....	51
Tabela 6.1: Características gerais dos pacientes envolvidos no estudo.....	55
Tabela 6.2: Distribuição dos pacientes com diferentes classes: 0 e 1 e para cada faixa temporal com o endpoint SLPD.....	55
Tabela 6.3: Distribuição dos pacientes com diferentes classes: 0 e 1 e para cada faixa temporal com o endpoint SG.....	55
Tabela 6.4: Distribuição dos pacientes com diferentes classes: 0 e 1 para o endpoint RECURR.....	56
Tabela 6.5: Características significativas para o endpoint SLPD com a faixa temporal Mediana, após os testes estatísticos.....	57
Tabela 6.6: Características significativas para o endpoint SLPD com a faixa temporal de 3 meses.....	59
Tabela 6.7: Características significativas para o endpoint SLPD com faixa temporal de 6 meses.....	62
Tabela 6.8: Características significativas para o endpoint SLPD com a faixa temporal 12 meses.....	63
Tabela 6.9: Características significativas para o endpoint SLPD com faixa temporal de 18 meses.....	65
Tabela 6.10: Características significativas para o endpoint SLPD com faixa temporal de 24 meses.....	67
Tabela 6.11: Características significativas para o endpoint SG com faixa temporal de 12 meses.....	70
Tabela 6.12: Características significativas para o endpoint SG com faixa temporal de 30 meses.....	71

Tabela 6.13: Sumário com as médias e os respectivos desvios padrão para as características significativas do endpoint SLPD em 6, 12, 18 e 24 meses.....	73
Tabela 6.14: Sumário com as médias e os respectivos desvios padrão para as características significativas do endpoint SG em 12 e 30 meses.....	73
Tabela 6.15: Características significativas para o endpoint RECURR.....	74
Tabela 6.16: Características significativas para o endpoint RECURR após a aplicação do algoritmo k-means e regressão logística.....	76

LISTA DE ABREVIATURAS

AM	-	Aprendizado de Máquina
AVC	-	Acidente Vascular Cerebral
CAD	, do inglês	<i>Computer-Aided Diadnosis</i> - Diagnóstico Auxiliado por Computadores
CTV	-	<i>Clinical Target Volume</i>
CQ	-	Computação Quântica
DT	, do termo em inglês	<i>Decision Tree</i> - Árvores de Decisão
GBM	-	Glioblastoma Multiforme
GTV	-	(<i>Gross Volume Tumor</i>)
IA	-	Inteligência Artificial
IRM	-	Imagens de Ressonância Magnética
IMRT	, do inglês	<i>Intensity Modulated Radiotherapy</i> - Radioterapia por Intensidade Modulada
KPS	-	Karnofsky <i>Performance Status</i>
LR	, do inglês	Logistic Regression - Regressão Logística
RECURR	-	Recorrência
RF	-	Random Forest
ROI	, do inglês	<i>Region of Interest</i> - Região de Interesse
SG	-	Sobrevida Global
SLPD	-	Sobrevida Livre da Progressão da Doença
SNC	-	Sistema Nervoso Central
TC	-	Tomografia Computadorizada
TMZ	-	Temozolamida
TPS	, do inglês	<i>Treatment Planning System</i> - Sistema de Planejamento
VMAT	, do inglês	Volumetric Modulated Arc Therapy - Radioterapia Arco Volumétrica

SUMÁRIO

1. INTRODUÇÃO.....	18
1.1 OBJETIVOS	18
2. GLIOBLASTOMA MULTIFORME (GBM).....	23
2.1 DEFINIÇÃO DE GBM	23
2.2 REPRESENTAÇÃO CLÍNICA DO GBM	24
2.3 TRATAMENTO	25
2.4 USO CLÍNICO DAS IMAGENS MÉDICAS.....	26
3. ANÁLISE DE TEXTURA	27
3.1 REVISÃO HISTÓRICA E DEFINIÇÃO DE RADIOMICS	27
3.2 FLUXOGRAMA DE RADIOMICS	28
3.3 LIMITAÇÕES DE RADIOMICS	29
4. APRENDIZADO DE MÁQUINA	31
4.1 ÁRVORE DE DECISÃO (DT) E RANDOM FORESTS (RF).....	36
4.2 REGRESSÃO LOGÍSTICA (LR)	42
4.3 ALGORITMO DE K-MEANS	43
5. METODOLOGIA	47
5.1 SELEÇÃO DE PACIENTES	47
5.2 AQUISIÇÃO DE BANCO DE DADOS	47
5.3 PRÉ PROCESSAMENTO DAS IMAGENS	48
5.4 EXTRAÇÃO DE CARACTERÍSTICAS POR RADIOMICS	49
5.5 AVALIAÇÃO DOS ENDPOINTS	50
5.6 AVALIAÇÃO DA DISTRIBUIÇÃO DE DOSE	52
5.7 ANÁLISE ESTATÍSTICA	52
6. RESULTADOS E DISCUSSÕES DOS DADOS	55
6.1 DISCUSSÕES GERAIS	80
7. CONCLUSÕES	83
8. REFERÊNCIAS BIBLIOGRÁFICAS	85

1. Introdução

Glioblastoma multiforme (GBM) é o tumor astrocítico cerebral mais comum em adultos, estatísticas americanas e europeias indicam uma incidência de 2 a 3 casos por 100 mil habitantes (Chaddad *et al.*, 2019). Seu tratamento padrão consiste da máxima ressecção cirúrgica possível, seguida de radio e quimioterapia, além de cursos de temozolamida adjuvantes (Stupp *et al.*, 2009). Trata-se de um dos tumores mais agressivos e letais em adultos, apresentando um mau prognóstico ao paciente, como um alto risco de metástases e uma sobrevida média de 14 meses (Adamson *et al.*, 2009). Suas características principais resumem-se em: baixa diferenciação histológica, alta atividade mitótica, heterogeneidade, áreas hemorrágicas e necróticas. Essa heterogeneidade dificulta seu tratamento, limitando, por exemplo, o uso de biópsias invasivas para caracterização do tecido.

O sistema TNM (descriptor do tamanho tumoral, seus linfonodos envolvidos e se há metástase distantes) de estadiamento tumoral, incorporado pelo Comitê Conjunto Americano de Estadiamento de Câncer (AJCC, do inglês *American Joint Committee on Cancer*), propõe uma base sólida para predição de sobrevida, escolha do tratamento adequado e estratificação de pacientes em triagens clínicas baseado em dados clínicos dos pacientes oncológicos. No entanto, atualmente, acredita-se que as imagens médicas possam ser a fonte de características quantitativas referentes às regiões anatômicas nelas presentes para auxílio a essas predições (Lambin *et al.*, 2013). Isso se aplica especialmente a casos de tumores heterogêneos, como o GBM, pela caracterização deste padrão de maneira não-invasiva. Assim, novas oportunidades com relação à caracterização tumoral através de biomarcadores devem surgir no campo da medicina.

No diagnóstico, os biomarcadores poderiam atuar na avaliação de risco, e também no estadiamento. Em seguida, poderiam ser utilizados no monitoramento da resposta tumoral a um determinado tratamento, podendo validar um modelo preditivo para uma terapia personalizada (Ludwig e Weinstein, 2005).

As imagens de ressonância magnética (IRM) são a modalidade de imagens usadas para diagnóstico e monitoramento do GBM. Acredita-se que elas contenham

informações sobre patologia, marcadores biológicos e também informações genômicas, que permitiriam seu uso para a extração de dados relacionados ao prognóstico da doença, o que pode permitir o auxílio à decisão clínica personalizada e de maior precisão (Aerts *et al.*, 2014; Lambin *et al.*, 2013). Assim, o tratamento personalizado mostra-se como alternativa para guiar o paciente a um melhor prognóstico, especialmente nos casos de GBM.

As imagens médicas são a figura principal desta modalidade de “medicina personalizada” não invasiva, porque, além de auxiliar no diagnóstico clínico, atuam como fonte de características quantitativas que podem ser extraídas através de algoritmos computacionais para futuramente correlacioná-las com os desfechos clínicos do paciente (do termo em inglês, *endpoints*). Exemplos desses *endpoints* são: sobrevida global (SG), sobrevida livre da progressão da doença (SLPD), e também com recorrência tumoral dentro do campo de tratamento (RECURR).

Essas características quantitativas são extraídas das imagens por algoritmos matemáticos e podem conter diversos tipos de informações como de formato, intensidade dos voxels, arranjo espacial dos voxels (textura), entre outros. Em 2010, o conjunto dessas características extraídas das imagens médicas por técnicas computacionais foram definidas pelo termo “Radiomics” (Gillies *et al.*, 2010), resultante da junção da palavra “radio” ao sufixo “omics”. O primeiro se refere ao diagnóstico por imagem, ciência que estuda imagens médicas através de radiação e o segundo, corresponde a um neologismo usado em biologia para se referir ao campo de estudo.

A extração de informações sobre o radiomics em IRM tem recebido interesse particular para segmentação de tecido e tumor em GBM (Assefa *et al.*, 2010; Bauer *et al.*, 2012), além da diferenciação do GBM de outros tipos de tumores cerebrais (Eliat *et al.*, 2012; Zacharaki *et al.*, 2009).

Atualmente, acredita-se ser possível aplicar técnicas de aprendizado de máquina (AM) para prever a SG dos pacientes com GBM (Zhou *et al.*, 2014) e também caracterizar aspectos clínicos do GBM (Yang *et al.*, 2015) através dessas

características, de forma a melhorar a qualidade de vida dos pacientes ao proporcionar uma medicina “personalizada” (Kumar *et al.*, 2012). Nesta linha, estudos mostram que é possível criar um classificador de prognóstico para pacientes de baixo e alto risco, bem como a SG baseados em características *Radiomics* extraídas de IRM (Chen *et al.*, 2019; Kickingeder *et al.*, 2016) através de algoritmos de aprendizado de máquina.

Outros estudos recentes indicam possíveis construções de modelos preditivos a partir da combinação de árvores de decisão (DT, do termo em inglês, *Decision Tree*) com características quantitativas para avaliar *endpoints* como SG e SLPD com diferentes abordagens de seleção de características e redução de dimensionalidade de dados (Chaddad *et al.*, 2019).

Com relação às sequências de IRM, Ingrisich *et al* (Ingrisich *et al.*, 2017) em sua revisão de estudos retrospectivos, mostra que é possível extrair informações prognósticas de GBM a partir de imagens ponderadas em T1 com contraste e com variações nos parâmetros de aquisição (resolução, espessura, tempo de eco, tempo de repetição e ângulo de rotação (Ingrisich *et al.*, 2017)) com *endpoint* sendo a SG dos pacientes. Em seu estudo com 66 pacientes, padronizou as imagens com uma resolução de 0,5 x 0,5 x 0,5 milímetros cúbicos, incluindo a região do tumor delineada manualmente. Após, 208 características foram extraídas, incluindo características com relação ao formato do tumor, estatísticas baseadas na intensidade dos pixels, características de Haralick, e pFTAS (do inglês *parameter-free threshold adjacency statistics*) através de algoritmos criados em Python. Para reduzir a dimensionalidade dos dados, fez-se uso do algoritmo *minimal depth variable selection* e então um algoritmo de *random forest* – com validação cruzada de 10 *folds* – foi utilizado como base para criação do modelo preditivo. Como resultado, encontrou 20 características com maior poder preditivo para implementar no modelo, e o índice de concordância foi de 67,7%.

De forma similar, (Bae *et al.*, 2018) criou modelos preditivos utilizando também *random forest*, mas adicionou uma nova classe de resposta como sendo a SLPD. 796 características (incluindo clínicas e genéticas) foram extraídas. A diferença está no método em que utilizou para redução da dimensionalidade dos dados: as

características foram dicotomizadas por um *threshold* e a função “cutp” do *software* matemático R (R *Foundation for Statistical Computing*, Vienna, Austria) que proporcionou a significância de cada uma. O modelo incluindo somente as características *Radiomics* obteve uma área sobre a curva (AUC) ROC de 65,2% para o *endpoint* SG e AUC de 59% para SLPD.

Chaddad *et al* (Chaddad *et al.*, 2018) criou uma metodologia na qual IRM ponderada em T1 e FLAIR de 40 pacientes foram usadas como referência para que as áreas do tumor ativo, necrose e edema fossem delineadas nas mesmas. A partir dessas imagens, extraiu-se características dessas regiões delineadas utilizando um filtro nas imagens baseado em Laplaciano de Gaussiano (do inglês *LoG filter*). Com essas características aplicou-se uma correlação de Spearman – o qual assimila-se com a correlação de Pearson, diferenciando-se no fato de que a correlação de Spearman pode ser aplicada para dados não lineares – para identificar quais características mostram-se mais correlacionadas (na ordem de 0,3 a 0,4) com os *endpoints* para o valor- $p < 0,05$. Ao avaliarem as características mais significantes, criou-se um modelo preditivo utilizando *Random Forests* para estimar a predição da SG e da SLPD. O parâmetro utilizado na validação do modelo foi uma validação cruzada com 5 *k-folds* e as observações foram aleatoriamente divididas entre conjunto de teste e validação. Ao final, obtiveram uma acurácia de 85% de predição da SG e 76,4% de acurácia da predição da SLPD para a região delineada como sendo o tumor ativo; 68% para a região de edema e 61,5% para a região de necrose. Da mesma forma, (Molina-García *et al.*, 2019) utilizou a correlação de Spearman, porém com uma ordem de grandeza de -0,7 e 0,7 para correlação.

Outra abordagem foi realizada por Chen *et al* (Chen *et al.*, 2019), por sua vez, utilizou um algoritmo de seleção de características baseada na mínima redundância acompanhada da máxima relevância das mesmas, para construção de três modelos preditivos para SG dos pacientes: um avaliando somente as características *Radiomics*; um avaliando somente características clínicas; e um último conjunto contendo todas as características. Esse algoritmo calcula a informação mútua de um conjunto de características com suas respectivas respostas e cria um *ranking* para as mais importantes. Nesse estudo 40 de 3824 características foram utilizadas para a implementação do modelo preditivo. O resultado final indicou que para o estudo

envolvendo *Radiomics* obteve-se uma AUC de 81,5% na validação; 67% para características clínicas, e para o último modelo AUC de 85,1%. O que indica que o uso do *radiomics* isoladamente apresenta um bom resultado quando comparado ao modelo mais complexo envolvendo *radiomics* e dados clínicos dos pacientes.

1.1 Objetivo

Neste contexto, este estudo tem como objetivo principal extrair informações quantitativas de imagens médicas de pacientes que foram submetidos ao tratamento radioterápico, com laudo de lesão de glioblastoma GBM através de processamento de imagens usando *radiomics*, a fim de se correlacionar estas características com determinados prognósticos da doença.

2 Glioblastoma Multiforme

2.1 Definição de Glioblastoma Multiforme

O Sistema Nervoso Central (SNC) é envolto de três camadas meningeais: *dura-mater*, *aracnoide* e *pia-mater*, e os tumores relacionados a este sítio anatômico são determinados pela histologia, grau e localização. Tratando-se especificamente de gliomas, a Organização Mundial da Saúde classifica-os em grau de 1 até 4, de acordo com o nível de malignidade. Gliomas de baixo grau possuem baixa proliferação e a cirurgia mostra-se eficaz para o tratamento, enquanto gliomas de alto grau – gliomas de grau 4 ou GBM – são os tipos de lesões do SNC mais agressivas e indiferenciadas, dificultando então o tratamento (JOVČEVSKA, KOČEVAR e KOMEL, 2013; Louis *et al.*, 2007). A tabela 2.1 descreve os gliomas com relação ao seu grau de malignidade.

Tabela 2.1: Definição do grau dos gliomas de acordo com a Organização Mundial da Saúde.

	Grau dos gliomas			
	I	II	III	IV
Descrição (Dhiman <i>et al.</i> , 2018)	Tumores com taxa de crescimento baixa, benignos e alto índice de sobrevida	Tumores com taxa de crescimento baixa porém com recorrência de alto grau, podendo ser malignos ou não	Tumores malignos com atividade mitótica e recorrência como gliomas de alto grau	Tumores com região necrótica, malignos, com alta taxa reprodutiva e alta agressividade

Fonte: Autoria própria

As principais características histopatológicas de tumores do tipo GBM são a alta atividade mitótica, alta proliferação vascular, heterogeneidade, hemorragias, regiões com edemas e regiões necróticas (Adamson *et al.*, 2009), e um exemplo de sua apresentação em IRM é notada na Figura 2.1.

Figura 2.1: A) Representação de GBM e B) Representação de edemas.

Fonte: (Steinfeld, 1992)

Esses tumores podem apresentar lesões metastáticas fora do SNC, devido à presença de células tronco cancerígenas que circulam pela vascularização tumoral. Essas condições indicam que esse determinado tipo celular possui a habilidade de migrar para outra região e assim iniciar uma etapa de proliferação celular, gerando novos tumores. (Humphries *et al.*, 2018; Navone *et al.*, 2019; Schaijik, van *et al.*, 2019).

2.2 Apresentação Clínica do GBM

Os sintomas do GBM podem se desenvolver rapidamente, o que pode ser confundido com um Acidente Vascular Cerebral – AVC (Omuro and DeAngelis, 2013). A apresentação clínica envolve os seguintes sintomas: deficit neuronal e cognitivo, náuseas, alta pressão intracranial e perda de memória (Hanif *et al.*, 2017).

Por sua vez, o diagnóstico pode ser feito através de biópsias e/ou exames de imagens médicas (TC – Tomografia Computadorizada ou IRM). Outros tipos de imagens podem auxiliar no diagnóstico, porém discutiremos essas duas modalidades, que são as mais usuais. Na imagem TC a lesão aparece hipointensa, enquanto nas IRM, dependendo da sequência de aquisição da imagem, pode aparecer hiperintensa, quando utilizamos um protocolo de imagem ponderada em T2 (Shukla *et al.*, 2017).

Normalmente, durante a aquisição da IRM, a administração do uso de um agente contrastante, como o gadolínio, é usada para maximizar o sinal obtido através da

aquisição de imagens e diferenciar a área necrótica da área que está envolta pelo tumor ativo.

2.3 Tratamento

O tratamento de tumores do tipo GBM é feito através da máxima remoção cirúrgica possível, e seguido de quimioterapia e/ou radioterapia concomitantes. Nos anos 2000 (Stupp *et al.*, 2009), um estudo randomizado da Organização Europeia de Pesquisa e Tratamento de Cancer (EORTC) e o Instituto Nacional de Câncer do Canadá (NCIC) com 573 pacientes diagnosticados com GBM e com idade entre 18 e 70 anos investigou o benefício do uso da substância Temozolamida (TMZ) como agente quimioterápico administrado concomitantemente com tratamento radioterápico. O resultado final obtido foi que, ao administrar TMZ junto à radioterapia aumentou a sobrevida global dos pacientes de 12 meses para 14 meses, quando comparado com um tratamento radioterápico sem quimioterapia (Cabrera *et al.*, 2016).

A radioterapia pode ser realizada com técnicas modernas e complexas, como Radioterapia com Intensidade Modulada (IMRT, do inglês *Intensity Modulated Radio Therapy*) ou Radioterapia Arco Volumétrica (VMAT, do inglês *Volumetric Arc Radio Therapy*), para garantir o tratamento de toda lesão e de minimizar a toxicidade em órgãos de risco como o tronco cerebral, medula e no próprio cérebro, minimizando a chance então de uma nova lesão radioinduzida. A entrega de dose para GBM totaliza de 59,4 Gy a 60 Gy com um fracionamento diário de 2 Gy/dia. Esse valor que está sujeito a alteração dependendo da performance do paciente (Cabrera *et al.*, 2016; Walker, Strike e Sheline, 1979).

Apesar das técnicas modernas de entrega de dose (IMRT e VMAT) e do uso de TMZ como agente quimioterápico serem eficazes no tratamento, a resposta tumoral a essas técnicas resulta em uma sobrevida global média de 9 meses para pacientes com KPS inferior a 70 e até 18 meses para pacientes com KPS superior ou igual a 70 (Maher *et al.*, 2001; Minniti *et al.*, 2008).

Diante deste mau prognóstico e também devido à heterogeneidade e à agressividade dos tumores GBM, é necessário conhecer o padrão dessa lesão para investigar

parâmetros genéticos que supostamente estão relacionados com a resposta tumoral ao tratamento, definindo então o conceito de medicina personalizada (O'Connor, 2017).

2.4 Uso clínico das imagens médicas

Os avanços da tecnologia envolvendo imagens médicas na oncológica proporcionou uma alta gama de ferramentas para diagnóstico tumoral, uma vez que os radio-oncologistas diagnosticam e classificam os tumores através do mapeamento dos mesmos nas imagens. Espera-se que as imagens médicas contenham, além de informações digitais, as informações sobre patologia, marcadores biológicos e também informações genômicas (Hou *et al.*, 2006) resumidas na Figura 2.2.

Figura 2.2: Aplicação simplificada de modelos preditivos para correlação com o prognóstico do padrão celular. À esquerda a representação histopatológica tumoral através da extração de características de imagens médicas (centro) e representação genômica do tumor, também pela extração de características.

Fonte: Autoria própria.

Os parâmetros citados acima podem variar, e as imagens médicas auxiliam na visualização dessas mudanças (O'Connor, 2017). Por exemplo, imagens médicas podem ser utilizadas para mapear regiões de tumores primários e metástases; avaliar resposta e controle tumoral ao tratamento indicado e detectar recorrência da lesão (Barrington *et al.*, 2014; Eisenhauer *et al.*, 2009).

Assim, uma alternativa para otimizar a escolha do tratamento seria a criação de um banco de dados que contenha características preditivas sobre o andamento do tratamento, como por exemplo SG média dos pacientes.

3. Análise de textura

3.1 Revisão histórica e definição de *Radiomics*

O uso de imagens médicas no auxílio diagnóstico levando em consideração características tumorais como descritores da imagem iniciou-se nos anos 1980, com o foco do diagnóstico auxiliado por computadores (CAD, do inglês *Computer-Aided Diagnosis*). O CAD utiliza um conjunto de características de imagens médicas – geralmente levando em consideração a distribuição espacial da intensidade dos pixels de uma região de interesse (ROI, do inglês *region of interest*) – para criar classificador a partir de inteligência artificial e assim classificar um determinado *endpoint* (Giger, Karssemeijer e Schnabel, 2013; Tang *et al.*, 2009).

Em oncologia clínica, esta técnica diferencia-se pelo fato de não focar em diagnóstico, e sim na resposta tumoral ao tratamento guiado pela equipe médica dos pacientes. A nova técnica é designada pelo nome *Radiomics*, cujo propósito é a extração de uma quantidade de características das imagens médicas (podendo ser superior ao CAD) para explorar possíveis relações entre o prognóstico do paciente que não podem ser analisados visualmente (Buvat, Orlhac e Soussan, 2015; Chicklore *et al.*, 2013; Kumar *et al.*, 2012; Lambin *et al.*, 2012).

As características extraídas utilizando-se "*Radiomics*" levam como base algoritmos matemáticos avançados e apresentam diversas categorias de extração de características quantitativas relativas à textura, e para cada uma destas categorias, existem subcategorias de características para serem analisadas (Amadasun e King, 1989; Davnall *et al.*, 2012; Dinstein, Shanmugam e Haralick, 1973; Gillies, Kinahan e Hricak, 2016; Lambin *et al.*, 2012; Scalco e Rizzo, 2017):

- I. Características Estatísticas de Primeira Ordem
- II. Características Baseadas em Geometria
- III. Matriz de Co-Ocorrência de Nível de Cinza
- IV. Zona de Tamanho de Nível de Cinza
- V. Matriz de Comprimento de Execução de Nível de Cinza
- VI. Matriz de Diferenciação de Nível de Cinza
- VII. Matriz Dependente do Nível de Cinza

Assim, como o *Radiomics* fornece a possibilidade de trabalharmos com diversas características a partir de imagens médicas, um grande potencial é esperado em aprimorar a qualidade de vida dos pacientes a partir de seu uso. Um estudo em 2017 (Avanzo, Stancanella e Naqa, El, 2017) revisou metodologias, estudos e resultados onde *Radiomics* é aplicado como ferramenta para criação de um modelo preditivo para diferentes tipos de *endpoints*, como mostra tabela 3.1 (Kickingreder *et al.*, 2016; Yang *et al.*, 2015):

Tabela 3.1: Resumo de metodologia e resultados para diferentes abordagens de aprendizado de máquina em pacientes com GBM.

Primeiro autor	Amostragem	Endpoint	IRM	Características	Modelo	Correlações
Kickingreder	119	Sobrevida Global	T1 pós contraste e FLAIR	Global, Textura e Formato	Principal Componente Supervisionado (PCA)	$p = 0.004$ na amostragem de validação
Yang	82	Sobrevida Média	T1, T2 e FLAIR	GLRLM, GLCM, Global, Textura, entre outras.	Random Forest	AUC = 0.69 para sobrevida média

Fonte: Adaptado de (Avanzo, Stancanella, & El Naqa, 2017).

3.2 Fluxograma do *Radiomics* para predição de *endpoints*

A prática de *Radiomics* para predição de determinado *endpoint* envolve etapas discretas, incluindo:

- i. Aquisição de imagens;
- ii. Identificação da ROI;
- iii. Segmentação da ROI (manual ou automaticamente);
- iv. Extração das características;
- v. Aplicação de Inteligência Artificial para criação de um modelo preditivo para um determinado *endpoint*.

Todas as etapas acima são ilustradas na Figura 3.1. A análise *Radiomics* inicia-se através da aquisição de imagens, nelas, a ROI envolvendo toda a região a ser analisada é delineada e/ou segmentada automaticamente. Em seguida, as características quantitativas são extraídas dessas ROIs para então serem aplicadas junto a um

algoritmo baseado em Inteligência Artificial para criar um modelo preditivo com base no *endpoint* desejado.

Figura 3.1: Representação do fluxo de trabalho da técnica Radiomics. (i) Aquisição de imagens; (ii) e (iii) Identificação e segmentação da ROI; (iv) extração das características Radiomics; (v) aplicação da Inteligência Artificial.

Fonte: Autoria própria.

3.3 Limitações do Radiomics

Quando imagens médicas são utilizadas na extração de características quantitativas para este tipo de análise, variações nos parâmetros de aquisição e/ou variações no protocolo clínico podem dificultar a investigação de efeitos biológicos nas imagens médicas a partir dessas características (Gillies, Kinahan e Hricak, 2016).

3.3.1 Aquisição de Imagens

(Galavis *et al.*, 2010) avaliou a variabilidade em 50 características em imagens PET (do inglês *positron emission tomography*) através de diferentes modos de aquisição de imagens e 40 das 50 características mostraram um desvio relativo de mais de 30%.

3.3.2 Segmentação

A acurácia relacionada à segmentação de uma lesão tumoral é totalmente afetada à variabilidade do observador (Rios Velazquez *et al.*, 2015). Além disso, o delineamento dessas estruturas gera um tempo excessivo do fluxo de trabalho de uma clínica oncológica. Por exemplo, (Lambin *et al.*, 2013) comparou a acurácia da segmentação semiautomática e manual do crescimento tumoral em imagens CT e tiveram como resultado uma melhor estabilidade na segmentação semiautomática.

3.3.3 Tamanho Tumoral e Heterogeneidade Intratumoral

Áreas necróticas e com edemas podem apresentar projeções de intensidade de pixel similares e podem ser confundidas quando as características Radiomics são extraídas. Como GBM apresenta uma alta heterogeneidade, com áreas necróticas centralizadas no centro da lesão e não há um padrão específico para a classificação tumoral, a análise Radiomics fica dificultada.

4. Aprendizado de Máquina

Assim como qualquer elemento, natural ou não, de difícil compreensão, temas como **Inteligência Artificial** (IA), **AM** (chamada de modo mais frequente em inglês de *Machine Learning*) ou **Computação Quântica** (CQ) causam diferentes reações em seres humanos, variando de medo e desconfiança a fascínio e admiração.

Um belíssimo exemplo de como a humanidade, em geral, reage inicialmente de modo cauteloso (ou furioso) foi a contestação de Nicolau Copérnico, que sugeria o Sol como elemento central (Lopes, 2014) ao redor dos quais os demais corpos do sistema solar orbitavam, contando ainda com uma interessante contribuição sugerindo, de modo ainda mais abrangente, um cosmo virtualmente infinito e com incontáveis mundos, muitos destes povoados e com os quais não temos contato.

Do mesmo modo que o embate geocentrismo/heliocentrismo gerou inflamados debates durante muitos anos, não foi diferente com Inteligência Artificial (*Machine Learning*). Embora não seja necessariamente simples ou fácil definir IA, pode-se compreender esse ramo da Computação como a capacidade de dotar um dispositivo computacional de elementos de software e dados para que uma dada tarefa seja realizada de modo similar – ou eventualmente melhor – aos humanos. Dentre os muitos modos possíveis, podemos focar nossa discussão em Aprendizado de Máquina, cuja delimitação conceitual é igualmente complexa.

Um modelo pode ser compreendido, em tal contexto, como uma relação matemática existente entre as diferentes variáveis consideradas no problema. Como exemplo de criação de um modelo a partir de dados, considere a Figura 4.1, construída usando um aplicativo qualquer para geração de gráficos, com base na tabela 4.1, descrita abaixo.

Tabela 4.1: Dados para exemplificar a construção de um modelo a partir de experimentos e porcentagem de redução de peso

Experimento	Redução de peso / %
1	5
2	8
3	11
4	14
5	18
6	20
7	25
8	26
9	29
10	31

Figura 4.1: Representação gráfica do modelo de redução de peso, com a regressão linear.

Usando um software para criação de gráficos e análise numérica básica como Microsoft Excel, podemos converter os dados brutos em informação visual, pois representamos os dados usando a relação existente entre eles, oriunda naturalmente da pesquisa que os gerou. As formas de representação visual da informação tendem a ser mais facilmente compreendidas do que a mesma na forma de tabelas numéricas. O gráfico representado na Figura 4.1 é chamado de diagrama ou gráfico de dispersão e permite a aplicação imediata de uma técnica estatística bem adequada ao problema.

Além da simples representação dos dados como pares de coordenadas cartesianas, o mesmo software efetuou uma operação estatística chamada regressão linear (Silva, 2016), que consiste na criação de uma equação representativa dessas coordenadas e a geração da chamada linha de tendências, especialmente útil quando se deseja efetuar uma extrapolação no gráfico, o que se confunde com uma operação de previsão, corroborando a capacidade preditiva desse modelo simplificado.

O pequeno exemplo citado acima ilustra a criação de um modelo que aprendeu a partir dos dados, e esta é uma das muitas definições disponíveis para Aprendizado de Máquina (Grus, 2015).

Existem diversos modelos disponíveis para explorar e analisar dados e construir, por exemplo, previsões a partir de dados novos. Podemos citar o aprendizado supervisionado (os dados estão identificados ou rotulados), aprendizado não-supervisionado (os dados não estão identificados ou rotulados), e online (o modelo recebe dados continuamente e ajusta-se de modo automático – não abordado neste documento) (Grus, 2015). A literatura contempla vários artigos de revisão relacionados com AM, apresentando a aplicação dos algoritmos em diferentes áreas do conhecimento humano (Caruana e Niculescu-Mizil, 2006).

Quando falamos a respeito de aprendizado supervisionado, podemos pensar em um modelo de ensino-aprendizagem muito utilizado por humanos, com a figura de um professor, que leciona um determinado tópico de conteúdo a seus estudantes e estes aprendem com base no modelo sugerido/apresentado pelo docente. Em termos de ciência de dados, seria o equivalente a apresentar ao algoritmo um certo conjunto de

números (*features*, na nomenclatura original em inglês) e depois um outro certo valor que pode representar alguma categoria na qual aquele conjunto de dados se insere. Como exemplo, os valores de massa, diâmetro e uma codificação adequada para textura são capazes de diferenciar uma maçã de uma laranja. Dessa forma, pode-se criar um classificador para efetuar essa distinção. Tal ideia foi usada neste trabalho e a abordagem escolhida será discutida mais adiante no documento.

O aprendizado não supervisionado envolve a análise de um conjunto de dados sem categorização prévia, onde os elementos podem ser, por exemplo, agrupados de acordo com suas características (ou *features*). Analogamente, pode-se comparar este modo de aprendizagem com aquele onde um grupo de alunos analisa um conjunto de informações e as categoriza segundo alguns critérios, sem a presença de um professor. Tal abordagem é amplamente usada em técnicas de agrupamento (ou *clustering*) e os chamados Mapas de Kohonen são um belo exemplo, conforme pode ser exemplificado na Figura 4.2 (Wehrens e Kruisselbrink, 2018).

Figura 4.2: Exemplo de classificação de tipos de pimenta.

A implementação do aprendizado de máquina consiste, resumidamente, em 5 etapas, como exemplifica a figura 4.3. A primeira delas é a aquisição dos dados (nesse trabalho representado como características Radiomics extraídas das IRM); a segunda é a mineração dos dados, ou seja, análise quantitativa das características para uma melhor representação das mesmas, etapa em que também são aplicadas técnicas de redução de dimensionalidade. A terceira, quarta e quinta etapas são, respectivamente: divisão dos dados entre conjunto de treinamento, teste e validação e aprimoramento do algoritmo escolhido. Porém, existem inúmeras formas de realizar a mineração dos dados e algoritmo de validação de modelo preditivo. E para este presente estudo, utilizamos teste t-Student, k-means, regressão logística e correlação de Spearman (usadas em diferentes abordagens) como mineração dos dados e algoritmos árvore de decisão (Vandenberghe *et al.*, 2017) e regressão logística como classificadores,

ambos algoritmos classificadores supervisionados. A seguir um detalhamento das abordagens utilizadas.

Figura 4.3: Fluxograma de aprendizado de máquina: 1) Aquisição dos dados; 2) Mineiração e manipulação dos mesmos; 3) Treinamento do modelo preditivo; 4 e 5) validação, teste e aprimoramento do mesmo.

4.1 Árvores de Decisão (DT) e Random Forests (RF)

Um aspecto essencial deste tipo de entidades está relacionado com sua representação abstrata. Nesta área do AM, a representação dessa estrutura de dados é inversa àquela encontrada em árvores reais: a raiz está no topo (figura 4.4).

Figura 4.4: Representação abstrata de uma árvore de decisões.

Fonte: Autoria própria.

De modo um pouco mais preciso e específico, a figura 4.5 apresenta um modo menos abstrato de visualização das DT (Breiman, L; Friedman, J. H.; Olsen, R. A.; Stone, 1984).

Figura 4.5: Representação mais precisa de uma DT.

Pela Figura 4.5, “Split n” indica o processo de partição dos dados em grupos menores.

Assim, de acordo com algum critério definido antes do processo de particionamento, o conjunto X foi dividido nos subconjuntos X_1 e X_2 que, por sua vez, foram divididos até que, novamente de acordo com algum critério previamente definido, fossem obtidos, por exemplo, os chamados nós X_{14} e X_{15} , aqui chamados de folhas ou terminais, pois não sofreram nenhuma divisão ou particionamento após sua criação.

Temos aqui o primeiro problema encontrado na classificação usando DT: como separar os dados para que os descendentes sejam mais “puros”, ou seja, mais específicos quando se considera um processo de separação e classificação. Outra maneira de compreender o problema seria algo como encontrar qual a proporção correta entre os dados do conjunto atual para obter os dois novos subconjuntos, considerando – claro – que o processo de divisão seja binário.

Uma ideia inicial a respeito das árvores de decisão pode ser vista na figura 4.6 que ilustra um pequeno trecho de um fluxograma, estratégia muito utilizada no projeto de algoritmos quando se exige uma tomada de decisão.

Figura 4.6: Exemplo de um trecho de fluxograma.

Fonte: Autoria própria

A figura 4.6 indica uma operação baseada na Lógica Booleana, explicitando alguns conceitos elementares como Exclusão Mútua e o Princípio do Terceiro Excluído. De modo simples, a Exclusão Mútua pode ser vista como uma ausência de dualidades, onde um dado argumento (ou valor/resultado) assume um ou outro valor dentro da faixa prevista, e não os dois valores possíveis. Assim, conforme mostrado na figura 4.6, o resultado do teste lógico pode ser apenas verdadeiro ou falso, nunca ambos. De modo igualmente simples, o Princípio do Terceiro Excluído garante que os valores mapeáveis sejam apenas os dois previstos na faixa de valores possíveis (zero ou um, neste caso) e nenhum outro. Tal rigor matemático garante a funcionalidade plena de aplicações baseadas na Lógica Booleana.

A Figura 4.7 ilustra uma tomada de decisão tendo por base apenas e tão somente a quantidade de patas de um ser vivo, visando classificá-lo como gato ou centopeia, neste exemplo bem específico.

Figura 4.7: Exemplo de decisão baseada em características: gato ou centopeia?

Fonte: Autoria própria.

Deste exemplo emerge uma questão crucial relacionada com as DT: qual a melhor pergunta a ser feita em cada etapa para separar o conjunto de dados em subconjuntos adequados. Uma pergunta má feita pode levar a um conjunto inadequadamente separado e, nitidamente, comprometer o processo de classificação devido ao chamado *overfitting*, ou sobreajuste, ou ainda supertreinamento. Não se trata de um bom cenário, uma vez que o processo todo pode estar comprometido, removendo parte da confiabilidade dos dados.

Os resultados provenientes de uma classificação usando DT podem ser apresentados de várias formas, incluindo resultados em forma textual/numérica ou na forma de gráficos, e as Figuras 4.8 e 4.9 ilustram essas possibilidades de exibição gráfica. Nota-

se, na Figura 4.8, que os dados usados nessa classificação foram agrupados em quatro conjuntos diferentes devido às respectivas características, compatível com os parâmetros iniciais de análise. Naturalmente, tais parâmetros foram definidos de acordo com a estratégia de análise seguida pelo analista que executou o procedimento de classificação, bem sólido e determinado, e poderia resultar, claramente, em dois conjuntos, dependendo dos rótulos (ou características) a serem classificados.

Figura 4.8: Resultados de classificação usando DT em forma de clusters.

Outro modo gráfico de visualizar os resultados da análise executada por um classificador baseado em DT pode ser visto na figura 4.9. Os dados exibidos estão fora do escopo deste trabalho e possuem caráter ilustrativo com relação à apresentação de resultados.

Figura 4.9: Apresentação gráfica dos resultados da classificação usando DT

Agora que sabemos como é o funcionamento de uma DT, apresentamos em seguida uma breve explicação e justificativa de porquê usar a Random Forest como classificador.

As DTs, como classificadores, são vulneráveis a *overfitting*, principalmente quando a mesma apresenta uma grande quantidade de ramificações. No caso de RF, um *ensemble* de DTs é utilizado, onde cada DT atua de maneira independente. Essa combinação de modelos torna a RF um classificador robusto e menos suscetível ao *overfitting*. Assim, cada DT vai gerar um resultado final de classificação dos dados, e o algoritmo de RF combina-os para gerar uma saída única.

4.2 Regressão Logística (LR)

Após a definição a respeito do tipo de aprendizado de máquina, o tipo de dado das características decide se o processo continua como classificação ou regressão: se o tipo de dado for contínuo, o mais adequado é uma regressão, mas quando o dado for discreto temos uma classificação.

Para classificar um dado paciente como objeto de recidiva de glioblastoma ou não, por exemplo, deve-se estimar as probabilidades relativas a cada opção de classificação com base nos valores associados a cada característica. Matematicamente, temos a probabilidade **P** do paciente **y** apresentar recidiva do glioblastoma (identificado como **gb**), dado o conjunto de dados **w**, é maior que a probabilidade **P** do paciente **y** não apresentar a mesma recidiva (identificado como **ngb**), como descrito na equação 4.1.

$$P(y=gb|w) \geq P(y=ngb|w) \quad (4.1)$$

Uma vez que os dados de saída são “binários” ou então “dicotômicos”, o método da regressão linear não é muito adequado para estimar essas probabilidades que necessitamos para classificar o paciente em relação à presença ou não de recidiva. Assim, a função sigmoide (equação 4.2), cuja equação é mostrada a seguir, é capaz de criar uma curva que permite um melhor ajuste e assim promover a classificação de modo mais adequado.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4.2)$$

O gráfico da função sigmoide pode ser usado, então para promover um cenário mais adequado de classificação, onde pode-se associar as probabilidades de pertinência² de cada paciente às regiões na curva ou próximas dela, conforma pode ser visto na Figura 4.10.

Figura 4.10: Representação gráfica da função sigmoide.

Embora seja bastante útil em experimentos de regressão logística, a função sigmoide pode apresentar resultados diferentes no mesmo cenário, principalmente quando existe correlação entre as variáveis (ou características) usadas na análise. A razão desse comportamento está no confundimento, onde a relação entre as variáveis pode levar a interpretações equivocadas, quando a adequada separação das variáveis, ou então a redução da quantidade de variáveis de análise, é realizada.

4.3 Algoritmo *k-means*

Sabe-se que o aprendizado de máquina pode ocorrer por ao menos três modos: supervisionado, não supervisionado e por reforço (este último fora do escopo deste trabalho). Analisando agora o aprendizado não supervisionado, no qual os dados de entrada não estão rotulados ou caracterizados, temos como um dos seus principais algoritmo aquele conhecido como K-means, ou algoritmo de agrupamento, também conhecido como algoritmo de *clustering*, desenvolvido em 1967 por James MacQueen (MacQueen e others, 1967).

A ideia central, considerando que os dados não são rotulados e não há possibilidade de classificação, é agrupar os dados nos chamados clusters, com base em alguma característica que os aproxima. Existem diversos cenários que podem ser contemplados com a análise viável no algoritmo, mas algumas limitações são conhecidas: (a) uso de dados apenas numéricos, dispostos como pontos em um plano cartesiano; (b) há um custo computacional relativamente alto, relacionado com o número total de elementos, de agrupamentos e de iterações.

De qualquer forma, o algoritmo funciona bem em muitos cenários e podemos considerar um exemplo simples que auxilia na compreensão do processo global.

A Figura 4.11 apresenta o resultado da aplicação do algoritmo k-means sobre um conjunto de 9 (nove) pontos dispostos em um dado plano cartesiano. Tais pontos podem representar qualquer tipo de dado e o que se sabe sobre eles restringe-se às coordenadas cartesianas.

Figura 4.11: Resultado do k-means, com 3 clusters ou agrupamentos.

O funcionamento do algoritmo k-means é simples e pode ser visto no diagrama a seguir.

Diagrama 4.1: Diagrama básico do algoritmo k-means.

Funcionamento básico do algoritmo k-means
Selecione k pontos como centroides iniciais
Repita
Calcule a distância de cada ponto a cada um dos centroides
Com base nas distâncias, forme os agrupamentos: um dado ponto pertence ao agrupamento do centroide de menor distância.
Calcule a distância média entre os pontos e o centroide de cada agrupamento
Reposicione os centroides nas médias calculadas
Até a convergência (nenhuma alteração nas posições dos centroides)

As distancias citadas no texto acima referem-se à distância euclidiana entre dois pontos dispostos em um plano cartesiano, conforme pode ser visto na Equação 4.4.

$$d_{a,b} = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2} \quad (4.3)$$

Um dos principais problemas envolvidos na aplicação do K-Means no aprendizado de máquina é a definição do número de agrupamentos quando se efetua a exploração de dados, ou mais especificamente, o valor de k. Uma das alternativas para estimar o melhor valor de k é usar um gráfico da soma do quadrado das distâncias (de cada ponto ao centroide) em função de um valor estimado de k, baseado no *Elbow Method* (Método do Cotovelo, em tradução livre) (Bholowalia e Kumar, 2014).

A Figura 4.12 exemplifica um gráfico onde se verifica a aplicação do *Elbow Method* para determinação do valor ótimo de k, na inflexão da curva onde se observa C3 no eixo horizontal.

Figura 4.12: Aplicação do Elbow Method no algoritmo K-means.

A escolha do valor ótimo de k, na inflexão é feita porque, quando aplicado o limite de C ao infinito, a distância entre os dados e os clusters são pequenas.

5. Metodologia

5.1 Seleção de pacientes

Esse estudo foi aprovado pelo Comitê de Ética da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto e do Hospital das Clínicas de Ribeirão Preto (HCFMRP-USP) (CAAE número 92848718.8.0000.5407), e incluiu 43 pacientes com lesão de glioblastoma multiforme que tiveram seguimento até seu óbito, sem distinção de gênero, etnia ou religião. Todos os pacientes foram submetidos à quimioterapia com tratamento concomitante e pelo menos um ciclo de tratamento adjuvante, além de radioterapia com dose entre 54 a 60 Gy, sem distinção de fracionamento (hipofracionamento ou convencional) ou técnica de tratamento (radioterapia conformacional ou de intensidade modulada). Foram excluídos pacientes menores de 18 anos e pacientes com ressecção total da lesão.

5.2 Aquisição do banco de dados

Para os 43 pacientes foram utilizados arquivos DICOM de IRM ponderada em T1 pós-contraste feita pré-tratamento e após ressecção parcial da lesão (somente para os pacientes que foram submetidos à cirurgia) adquiridas pelos tomógrafos Siemens Achieva com campo magnético estático B0 de 1.5 e 3T, através do sistema de gerenciamento DICOM K-PACS, utilizado como banco de dados no Hospital das Clínicas de Ribeirão Preto.

Essas imagens foram importadas no sistema de planejamento (TPS, do termo em inglês *Treatment Planning System*) Eclipse (Varian Medical Systems) para que, com o auxílio de um radio-oncologista, o conjunto de estruturas contendo os alvos de tratamento GTV (*Gross Tumor Volume*) e CTV (*Clinical Target Volume*) fosse delineado, como mostrado na Figura 5.1. O GTV foi delineado em pacientes que não passaram por cirurgia antes do tratamento, e o CTV foi delineado em pacientes que tiveram ressecção parcial da lesão antes do tratamento radioterápico (exemplo na Figura 5.1). Nessas imagens, em uma etapa futura, as características quantitativas das estruturas vão ser extraídas a partir da técnica de *Radiomics*. Apenas um médico radio-oncologista foi o responsável por todos os delineamentos, o que nos permitiu reduzir as variações nos contornos das estruturas avaliadas.

Figura 5.1: Representação do delineamento das estruturas GTV e CTV em uma IRM. (a) IRM sem conjunto de estruturas; (b) IRM com conjunto de estruturas..

Fonte: Autoria própria

Além disso, para os mesmos 43 pacientes, foram exportados dos TPSs XiO (Elekta Medical Systems) e Eclipse (Varian Medical Systems) os arquivos DICOM RT de cada planejamento, contendo os arquivos DICOM da CT de planejamento, o arquivo DICOM RT Plan – responsável por representar a composição dos campos e algoritmo de cálculo de dose do planejamento–, o arquivo o arquivo DICOM RT Structure – responsável pelo armazenamento das estruturas delineadas e o arquivo DICOM RT Dose – responsável por copiar a matriz de dose dos sistemas de planejamento e correlacionar com os voxels da CT. Esses arquivos do planejamento do tratamento foram fundidos com as IRM e permitiram determinar a porcentagem de dose de radiação recebida em 95% do volume avaliado, um parâmetro clínico de avaliação do plano de tratamento.

5.3 Pré-processamento das imagens

Com o conjunto de estruturas delineado junto a IRM, as mesmas foram importadas ao software 3D Slicer – um software gratuito e aberto para processamento de imagens (www.slicer.org). Este software foi desenvolvido pelo Centro Nacional de Terapia Guiada por Imagem (do inglês NCIGT: *National Center for Image Guided Therapy*)

– entre outros institutos, e faz uso de plugins criados através da linguagem Python para aplicar diversas análises em um mesmo conjunto de imagens.

Ao importar as imagens dos pacientes, observou-se que as dimensões das IRM em comprimento e largura eram variadas. Visto que a técnica Radiomics é sensível à dimensão da imagem (Yip e Aerts, 2016), optou-se por padronizar todas as imagens para uma dimensão única de 240x240 pixels (resolução que era mais frequente). A espessura de cada imagem era dependente do protocolo clínico utilizado na aquisição da mesma nos tomógrafos de ressonância.

Esse processo de ajuste do tamanho da imagem foi feito no próprio 3D Slicer, em um módulo de deformação de imagens chamado “*Resize Image (BRAINS)*” no qual é necessário definir uma imagem Input, Output e o fator de escala, responsável por expandir ou reduzir as dimensões simetricamente, e também há um módulo chamado “Volumes” em que é possível visualizar as dimensões das imagens.

Dessa forma, a imagem Input foi escolhida, para todos os pacientes, como sendo a imagem de referência T1 pós contraste, e àquelas que não possuíam dimensões 240x240 pixels, foi escolhido um fator de escala apropriado para que o Output contivesse uma imagem com dimensão 240x240 padronizada, de acordo com a equação 5.1:

$$Output = \frac{1}{F} D_{(x,y)} \quad (5.1)$$

Onde $D_{(x,y)}$ representa a dimensão original da imagem, e F é o fator de escala da imagem Output. Dessa forma, para imagens com dimensões maiores que 240x240 pixels um $F > 1$ foi utilizado justamente para reduzi-las; enquanto para imagens com dimensões menores que 240x240 um $F < 1$ foi utilizado para expandi-las.

5.4 Extração de características pelo uso de Radiomics

O plugin Radiomics foi instalado no 3D Slicer. Nesse plugin foi escolhida a imagem de referência pré-processada, quando necessário, e a respectiva segmentação do GTV,

para que 105 características quantitativas fossem extraídas. (Lambin, 2016) descreve as classes de características utilizadas no Radiomics:

- i. Estatística de Primeira Ordem: Descreve a distribuição da intensidade do voxel dentro de uma região da imagem definida por uma máscara utilizada para métricas básicas;
- ii. Características de Forma 3D: Classe de descritores tridimensionais de uma ROI. As características dessa classe independem da distribuição de intensidade do pixel;
- iii. Matriz de Coocorrência de Nível de Cinza: Descreve uma função de probabilidade de uma ROI definida por uma máscara. O resultado final é uma matriz cujos elementos representam os valores de pixel e suas respectivas ocorrências;
- iv. Matriz de Região de Escala de Cinza: Quantifica zonas de escala de cinza em uma imagem. Uma zona de escala é definida como o número de voxels conectados com o mesmo valor de intensidade de pixel;
- v. Matriz de comprimento de Nível de Cinza: Quantifica execuções de nível de cinza que são definidas como o comprimento em número de pixels, de pixels consecutivos que possuem o mesmo valor em escala de cinza;
- vi. Matriz de Diferença de Escala de Cinza em Vizinhança: Quantifica a diferença entre o valor de intensidade de um pixel com a média do valor em sua vizinhança dentro de uma distância conhecida;
- vii. Matriz de Dependência de Nível de Cinza: Quantifica a dependência de escala de cinza para um pixel. Por sua vez, a dependência é definida como o número de voxels conectados dentro de uma distância conhecida cuja diferença entre um voxel e seu consecutivo é menor que um valor conhecido, também.

5.5 Avaliação dos *endpoints*: Recidiva (RECURR), Sobrevida Global (SG) e Sobrevida Livre da Progressão da Doença (SLPD)

Para cada paciente avaliou-se três endpoints junto ao prontuário clínico:

- i. Sobrevida Livre de Progressão da Doença (SLPD): Tempo estimado em meses em que os pacientes estiveram livre da doença até o seu reaparecimento, quando aplicado;

- ii. Sobrevida Global (SG): Tempo estimado em meses em que os pacientes estiveram vivos após o diagnóstico e/ou cirurgia de recessão parcial da lesão.
- iii. Recidiva (RECURR): reaparecimento da doença dentro do campo de tratamento, sem considerações sobre o tempo envolvido. Os pacientes que tiveram evento de recidiva nessas condições foram atribuídos com o valor 1; enquanto para os pacientes que não tiveram esse evento foram caracterizados com número 0;

Após a avaliação, como uma primeira estimativa temporal para determinado desfecho clínico, obteve-se a mediana da distribuição da SLPD e SG, porque assim garantimos que os pacientes estejam homogeneamente distribuídos em 50% com SLPD e SG superiores à mediana e 50% inferiores à mediana. Dessa forma, os pacientes cujos valores encontravam-se abaixo da mediana receberam valor 0, enquanto os pacientes cujos valores encontravam-se superiores à mediana receberam valor 1. Além disso, separou-se os pacientes em grupos temporais para os grupos SG e SLPD de acordo com a ocorrência, para avaliar as características mais significativas em uma distribuição não homogênea dos dados. A tabela 5.1 sintetiza todos os *endpoints* utilizados neste estudo. Os pacientes que apresentavam SLPD ou SG menor do que o valor temporal empregado receberam valor 1 e para os pacientes com SLPD ou SG acima desses valor receberam valor 0.

Tabela 5.1: Representação dos endpoints.

Endpoint	Tempo em meses para divisão dos grupos
SLPD MEDIANA	23
SLPD3	3
SLPD6	6
SLPD12	12
SLPD18	18
SLPD24	24
SG MEDIANA	16
SG3	3
SG6	6

SG12	12
SG18	18
SG24	24
SG30	30
SG36	36
RECURR	Não se aplica

5.6 Avaliação da distribuição de dose nos volumes alvo GTV e CTV

Com os arquivos DICOM RT dos pacientes, foi possível importá-los no TPS Monaco (Elekta Medical Systems) para avaliação analítica da dose relativa recebida em 95% do volume avaliado (D95%), como visto na Figura 5.2. Este dado seria uma métrica de qualidade no tratamento realizado nos pacientes, uma vez que pacientes tratados de maneira inadequada há uma chance maior de recidiva, ou redução da SLPD e SG. A escolha de utilizar a análise a partir da dose e volume relativos teve como objetivo principal eliminar a dependência da dose de prescrição.

Figura 5.2: Exemplo de avaliação da D95% no DVH para GTV/CTV.

5.7 Análise estatística

Após a extração de todos os dados citados acima, várias tabelas foram montadas, uma para cada endpoint. Cada tabela contém 43 linhas, representadas pelos pacientes, e

107 colunas, a primeira com um número de identificação do paciente e as 106 demais com preditores (105 características Radiomics e a D95%).

Com todos esses dados, buscamos reduzir o número de características que se relacionem a cada endpoint. Para isso para cada endpoint estudado, separamos os pacientes em duas tabelas, uma com os pacientes com endpoint 1 e outra com endpoint 0 e posteriormente fizemos a comparação das colunas com o teste t de Student para avaliar se existe uma diferença significativa estatisticamente é encontrada. Para realizar esse teste é necessário que as colunas comparadas tenham uma distribuição normal (pré-requisito para o teste) e, para as características com essa distribuição, aplicamos o teste de hipótese, onde a hipótese nula assume que não existe diferença entre as médias de uma mesma característica e a hipótese alternativa mostra que existe.

Toda a análise foi feita em Python. Para as análises estatísticas utilizou-se a biblioteca Scipy com o modulo stats (SCIPY, 2019). Primeiramente para importar as tabelas em CSV, utilizou-se a biblioteca Pandas (PANDAS, 2019). Após a importação, removeu-se a coluna ID que tem código de identificação do paciente, fez-se uma transformação logarítmica seguida da normalização dos dados pelo valor máximo. Em seguida, para cada endpoint, dividiu-se a tabela e duas novas tabelas, dependendo do seu endpoint. Aplicou-se o teste de Shapiro-Wilk, para avaliação da normalidade de cada característica e em seguida, para as de distribuição normal, aplicou-se o test-t de Student para normais a fim de avaliar se eram estatisticamente diferentes para um nível de significância de 5% ($p < 0,05$).

Para os pacientes com *endpoint* RECURR outras duas abordagens foram realizadas para a mineração dos dados:

- i. A primeira se baseou em uma clusterização a partir do algoritmo *k-means* com $k = 2$ clusters (Etapa 1). Com isso, algumas características foram selecionadas, mas como ainda eram muitas para serem utilizadas em um modelo preditivo sem sobreajuste (*overfitting*) dos dados, aplicou-se na sequência uma regressão logística para a seleção final das características significativas. Os dados empregados nesta abordagem foram normalizados pelo desvio padrão,

foram considerados significativos os resultados com $p < 0,05$ e o software utilizado para esta análise foi o IBM SPSS, versão 25. Com essas características criou-se no *Classification Learner* em Matlab® um modelo preditivo utilizando LR com $k\text{-fold} = 5$ para validação cruzada.

- ii. A segunda reproduziu a metodologia proposta no artigo (Chaddad *et al.*, 2018), o qual utiliza o ranking de *Spearman* – responsável por classificar as correlações entre características e seus respectivos desfechos clínicos (Zar, 1972) como método de redução de dimensionalidade e também utiliza o algoritmo de *Random Forests* com validação cruzada utilizando 5 $k\text{-folds}$ para construção do modelo preditivo. Ao aplicar o teste de *Spearman*, todas as características com correlação da ordem de grandeza entre 0,3 e 0,4 e com respectivo valor- $p < 0,05$ foram utilizadas como características mais significativas para construção do modelo preditivo. Toda esta análise foi realizada no *Classification Learner* em Matlab®.

6 Resultados e discussões de dados

Um total de 43 pacientes foram estudados (tabela 6.1), 60% dos pacientes são do sexo masculino e 40% do sexo feminino, as idades no momento do diagnóstico variaram entre 34 e 85 anos e 26 deles apresentaram recidiva da doença no campo de tratamento. A SG média foi de $21,1 \pm 14,4$ meses e SLPD média de $13,5 \pm 13,3$ meses, após receberem uma D95% média de $92,1 \pm 5,4\%$.

Tabela 6.1: Características gerais dos pacientes envolvidos no estudo.

Característica	Valores
Idade no Diagnóstico (Media/Desvio Padrão/Mínimo/Máximo em anos)	58,9/11,2/34/85
Gênero (M/F)	26/17
Recorrência (sim/não)	26/17
SG (Media/Desvio Padrão/Mínimo/Máximo em meses)	21,1/14,4/1/56
SLPD (Media/Desvio Padrão/Mínimo/Máximo em meses)	13,5/13,3/0/52
D(95% V) (Media/Desvio Padrão/Mínimo/Máximo) (%)	92,1/5,4/67/100

Para a avaliação dos diversos endpoints propostos, a distribuição dos pacientes nos diferentes grupos é representada pelas tabelas 6.2, 6.3 e 6.4, lembrando que a classe 1 corresponde a pacientes que recidivaram no período de tempo estipulado e a classe 0 corresponde a pacientes que não recidivaram neste tempo.

Tabela 6.2: Distribuição dos pacientes com diferentes classes: 0 e 1 e para cada faixa temporal com o endpoint SLPD.

CLASSE	Número de pacientes					
	SLPD MEDIANA	SLPD3	SLPD6	SLPD12	SLPD18	SLPD24
1	21	6	17	30	33	35
0	22	37	26	13	10	8

Tabela 6.3: Distribuição dos pacientes com diferentes classes: 0 e 1 e para cada faixa temporal com o endpoint SG.

CLASSE	Número de pacientes							
	SG MEDIANA	SG3	SG6	SG12	SG18	SG24	SG30	SG36
1	22	1	6	13	25	29	33	35
0	21	42	37	30	18	14	10	8

Tabela 6.4: Distribuição dos pacientes com diferentes classes: 0 e 1 para o endpoint RECURR.

CLASSE	Número de pacientes
	RECURR
1	26
0	17

A seguir serão apresentados os resultados obtidos para cada *endpoint* avaliado. São apresentadas as características com distribuição normal (em números) obtidas após o teste de Shapiro-Wilk e em seguida, as características com diferenças significativas entre o *endpoint* 0 e 1 no teste t de *Student*.

- **SLPD**

Para a SLPD MEDIANA das 105 características totais extraídas, 33 possuem distribuição gaussiana, ou seja, poderiam ser usadas no teste-t. Após a realização do teste de hipótese, ao se avaliar o p-valor, somente 7 delas mostram-se significativas (Tabela 6.5), o que significa que elas possuem médias diferentes, podendo ser utilizadas em um futuro modelo preditivo para separação dos grupos de pacientes que recidivaram ou não dentro da margem temporal mediana (23 meses).

Table 6.5: Características significativas para o endpoint SLPD com a faixa temporal Mediana, após os testes estatísticos.

SLPD MEDIANA	
Características com distribuição gaussiana	33
Características com diferença significativa entre os dois grupos	p-valor
<i>GrayLevelNonUniformity</i>	0,048
<i>MajorAxis</i>	0,004
<i>Maximum2DDiameterColumn</i>	0,019
<i>Maximum2DDiameterSlice</i>	0,006
<i>Maximum3DDiameter</i>	0,011
<i>Strength</i>	0,019
<i>Volume</i>	0,028

A Figura 6.1 apresenta o gráfico da dispersão em pares das 7 características selecionadas para seus respectivos grupos: azul para pacientes que não recidivaram em 23 meses (valor 0), e laranja para pacientes que recidivaram (valor 1). A diagonal principal corresponde distribuição dos valores da característica encontrados para os dois grupos e permite a verificação da sobreposição dos dados analisados. A última coluna e última linha mostram gráficos com o valor de cada característica para os *endpoints* (0 ou 1) e novamente verifica-se a dificuldade de se determinar um limiar que possa ser aplicado para separação dos grupos. Assim, com essas 7 características, possíveis construções de modelos preditivos podem ser prejudicadas, uma vez que não foi possível separar os valores obtidos para as características para os dois *endpoints*.

Figura 6.1: Gráfico de dispersão da característica SLPD com a faixa temporal Mediana.

Para os pacientes que recidivaram em até 3 meses, foram obtidas 43 características com distribuição gaussiana e apenas uma significativa para o teste de hipótese realizado, a *Kurtosis* (Tabela 6.6). Pelo gráfico da dispersão em pares (Figura 6.2), vemos que a região de valor 0,4 da *Kurtosis*, em escala logarítmica e normalizada mostra-se como um bom limite de separação dos grupos de pacientes que recidivaram em até 3 meses. Este valor limite encontrado equivale a uma *Kurtosis* de 2,717 na escala dos valores de saída direta do *radiomics* no *3D Slicer*. Isso indica que os

pacientes cujo valor de *Kurtosis* esteja abaixo desse limiar, podem apresentar recidiva da lesão em até três meses. Assim, essa característica mostra-se um indicador relevante para identificação dos tumores com característica agressiva e consequentemente mal prognóstico.

Tabela 6.6: Características significativas para o endpoint SLPD com a faixa temporal de 3 meses.

SLPD 3	
Características com distribuição gaussiana	43
Características com diferença significativa entre os dois grupos	p-valor
<i>Kurtosis</i>	0,016

Figura 6.2: Gráfico de dispersão para a característica SLPD com a faixa temporal de 3 meses e o limiar da *Kurtosis* em 0,4 (escala logarítmica).

Dado que a variável *Kurtosis* é um indicativo de possível separação dos grupos de paciente, aplicando-se um modelo preditor da recidiva em até 3 meses, baseado diretamente neste limiar de *kurtosis* encontrado aos pacientes incluídos neste estudo, obtivemos a matriz de confusão (figura 6.3). Podemos observar que esse modelo acerta 100% dos casos de recidiva, no entanto, para os casos de não recidiva o acerto é de 65%. A acurácia do modelo, dada pela porcentagem acertos (%verdadeiro positivo + %verdadeiro negativo), é de 70%.

Figura 6.3: Matriz de confusão para o endpoint SLPD com a faixa temporal de 3 meses, feita a partir da análise estatística para estimar a acurácia dessa separação.

Aplicou-se também o algoritmo de Regressão Logística com validação cruzada utilizando 5 folds, a variável *Kurtosis* como preditor e o endpoint SLPD em 3 meses como variável de resposta, para estimar a acurácia desse preditor e comparar com a matriz de confusão. Uma vez que este algoritmo apresenta boa resposta para dados contínuos e com saída binária, como no caso desse estudo, obteve-se a curva ROC e a matriz de confusão representadas na Figura 6.4.

Figura 6.4: Curva ROC e matriz de confusão para o modelo preditivo baseado em LR para SLPD em 3 meses.

Esse modelo proporcionou uma AUC de 75% e a matriz de confusão com uma acurácia de 88,4%. A acurácia do modelo baseado em LR é maior quando comparada com a matriz de confusão exemplificada na figura 6.3. Diante disso, nota-se que a variável *Kurtosis* junto com o modelo preditivo baseado em LR apresentam bons resultados para classificação de pacientes com mau prognóstico de GBM – recidiva de até 3 meses, nessa análise.

Por definição, *Kurtosis* caracteriza o achatamento da curva de distribuição de probabilidade, e seu valor pode representar uma distribuição concentrada na média, caso seu valor seja positivo; se possuir valor negativo, a concentração está distribuída possui maior concentração nas caudas. Agora, se seu valor for nulo, tem a mesma representação de uma distribuição normal (Lambin, 2016).

Para os intervalos temporais de SLPD de 6, 12, 18 e 24 meses, obteve-se respectivamente, as tabelas 6.7, 6.8, 6.9 e 6.10 e as figuras 6.5, 6.6, 6.7 e 6.8. As tabelas apresentam número de características com distribuição gaussiana e a indicação das características com significância estatística para cada conjunto de dados. Já as figuras apresentam a distribuição em pares dos valores obtidos.

Tabela 6.7: Características significativas para o endpoint SLPD com faixa temporal de 6 meses.

SLPD6	
Características com distribuição gaussiana	30
Características com diferença significativa entre os dois grupos	p-valor
MajorAxis	0,043

Figura 6.5: Gráfico de dispersão para o endpoint SLPD com faixa temporal de 6 meses.

Tabela 6.8: Características significativas para o endpoint SLPD com a faixa temporal 12 meses.

SLPD12	
Características com distribuição gaussiana	22
Características com diferença significativa entre os dois grupos	p-valor
<i>MajorAxis</i>	0,013
<i>Maximum2DDiameterSlice</i>	0,023
<i>MinorAxis</i>	0,039

Figura 6.6: Gráfico de dispersão para o endpoint SLPD com faixa temporal de 12 meses.

Tabela 6.9: Características significativas para o endpoint SLPD com faixa temporal de 18 meses.

SLPD18	
Características com distribuição gaussiana	38
Características com diferença significativa entre os dois grupos	p-valor
MajorAxis	0,007
<i>Maximum2DDiameterColumn</i>	0,028
<i>Maximum2DDiameterRow</i>	0,024
<i>Maximum2DDiameterSlice</i>	0,021
<i>Maximum3DDiameter</i>	0,011
<i>MinorAxis</i>	0,016

Figura 6.7: Gráfico de dispersão para o endpoint SLPD com faixa temporal de 18 meses.

Tabela 6.10: Características significativas para o endpoint SLPD com faixa temporal de 24 meses.

SLPD24	
Características com distribuição gaussiana	40
Características com diferença significativa entre os dois grupos	p-valor
<i>MajorAxis</i>	0,006
<i>Maximum2DDiameterRow</i>	0,008
<i>Maximum2DDiameterSlice</i>	0,009
<i>Maximum3DDiameter</i>	0,007
<i>MinorAxis</i>	0,021

Figura 6.8: Gráfico de dispersão para o endpoint SLPD com faixa temporal de 24 meses.

Para esses quatro intervalos temporais, a sobreposição dos valores encontrados para as características nos dois grupos ocorre novamente, como foi visto para a mediana de SLPD. Apesar de haver características significativas para esses endpoints, nenhum deles possui um limite pelo qual a separação de dados é possível. Mas, ao olharmos as características significativas nesses quatro grupos temporais, incluindo também a mediana de SLPD, é possível ver que a característica *Major Axis* aparece em todos os grupos. De acordo com (Lambin, 2016), essa característica é calculada por:

$$major\ axis = \sqrt[4]{\lambda_{major}}$$

Onde λ_{major} representa a maior componente do eixo da ROI – nesse caso, a maior componente 3D do GTV/CTV.

Dessa forma, para tumores GBM menos agressivos, que recidivam em um tempo maior que 3 meses, essa característica pode fornecer informações relevantes quanto à SLPD em até 24 meses. Um valor pequeno pode indicar uma menor agressividade da lesão, e consequentemente uma melhor qualidade de vida do paciente.

- **SG**

Para o *endpoint* SG, obteve-se para os grupos temporais separados pela mediana, 3, 6, 18, 24 e 36 meses, respectivamente, os seguintes números de características com distribuição gaussianas: 58, 0, 39, 54, 46 e 38. Somente um paciente apresentou SG em até 3 meses, o que não permite a análise estatística proposta para esta escala temporal, por isso a representação como zero. No entanto, após a realização do teste-t, nenhuma característica significativa capaz de separar os valores médios dos dois grupos foi encontrada. Portanto, essa metodologia empregada não foi capaz de separar os grupos de pacientes de forma adequada nestas escalas temporais.

Para SG de 12, observamos que 33 características com distribuição normal foram encontradas e dentre elas, 3 características apresentaram significância estatística para diferenciação das médias dos dois grupo (tabela 6.11). A distribuição em pares dos valores obtidos para as 3 características significativas (figura 6.9), indicam novamente uma sobreposição das distribuições de valores. No entanto, foi possível identificar um paciente com comportamento de *outlier* dentre os pacientes classificados como 1, pois através de uma análise da tabela de dados deste estudo os valores das 3 características dos pontos indicados correspondem ao mesmo paciente (destaque em vermelho na Figura 6.9).

Tabela 6.11: Características significativas para o endpoint SG com faixa temporal de 12 meses.

SG12	
Características com distribuição gaussiana	33
Características com diferença significativa entre os dois grupos	p-valor
<i>MajorAxis</i>	0,012
<i>Maximum2DDiameterSlice</i>	0,041
<i>Maximum3DDiameter</i>	0,034

Figura 6.9: Gráfico de dispersão para o endpoint SG com faixa temporal de 12 meses. O destaque em vermelho representa um paciente classificado como outlier.

Por fim, para SG de 30 meses, 42 características com distribuição normal foram encontradas, resultando em 5 características com significância estatística para diferenciação das médias dos dois grupo (tabela 6.12). A distribuição em pares dos valores obtidos para essas características significativas (figura 6.10) indicam novamente uma sobreposição das distribuições de valores, que impossibilita uma separação simples dos dados, embora suas médias sejam distintas.

Tabela 6.12: Características significativas para o endpoint SG com faixa temporal de 30 meses.

SG30	
Características com distribuição gaussiana	42
Características com diferença significativa entre os dois grupos	p-valor
<i>MajorAxis</i>	0,015
<i>Maximum2DDiameterRow</i>	0,016
<i>Maximum2DDiameterSlice</i>	0,023
<i>Maximum3DDiameter</i>	0,012
<i>MinorAxis</i>	0,025

Figura 6.10: Gráfico de dispersão para o endpoint SG com faixa temporal de 30 meses.

Percebe-se que as características significativas presentes nos grupos SG de 12 e 30 meses aparecem também para o grupo SLPD (vide tabelas 6.13 e 6.14, respectivamente), que por sua vez fazem parte do conjunto de caracterização do formato tumoral dentro do *Radiomics*. Dessa forma, a variação das médias dessas características foi feita ao longo de cada grupo temporal para os dois grupos de pacientes, e está representada nas tabelas 6.13 e 6.14.

Tabela 6.13: Sumário com as médias e os respectivos desvios padrão para as características significativas do endpoint SLPD em 6, 12, 18 e 24 meses.

Características	SLPD6		SLPD12		SLPD18		SLPD24	
	Valor 0	Valor 1	Valor 0	Valor 1	Valor 0	Valor 1	Valor 0	Valor 1
Major Axis (Media/Desvio Padrão)	52,1/18,9	67,6/26,9	45,6/15,3	63,7/24,4	43,2/16,5	62,8/23,5	41,6/18,3	62,1/23,9
Maximum2DDiameterSlice (Media/Desvio Padrão)	-	-	49,6/19,8	66,6/22,9	47,8/20,8	65,6/22,5	45,3/22,7	65,2/21,9
Minor Axis (Media/Desvio Padrão)	-	-	34,4/15	43,3/13,6	32,2/15,1	43,1/13,5	31,7/16,1	42,7/13,5
Maximum3DDiameter (Media/Desvio Padrão)	-	-	-	-	52,2/21,4	72,9/21,2	50,2/23,8	72,2/23,6

Tabela 6.14: Sumário com as médias e os respectivos desvios padrão para as características significativas do endpoint SG em 12 e 30 meses.

Características	SG12		SG30	
	Valor 0	Valor 1	Valor 0	Valor 1
Major Axis (Media/Desvio Padrão)	51,4/15,9	73,9/30,5	44,7/17,9	62,3/23,5
Maximum2DDiameterSlice (Media/Desvio Padrão)	56,1/20,4	73,9/24,9	48,3/21,3	65,5/22,5
Maximum3DDiameter (Media/Desvio Padrão)	61,8/20,2	82,6/29,4	52,6/21,9	72,8/21,2

Diante do exposto na tabela 6.13 (SLPD), percebe-se uma diminuição de 20,2% da variável *Major Axis*, que representa o maior eixo tumoral, para pacientes que não recidivaram de 6 até 24 meses. Este resultado é plausível uma vez que quanto menor o tumor, menor sua agressividade. Para os pacientes que recidivaram em até 24 meses, a variação existe, porém com um valor menor: 8%. Quando passamos para a tabela 6.14 (SG), para o *Major Axis* a variação é de 13% para pacientes que tiveram sobrevida global maior que 12 e 30 meses. A mesma justificativa é válida para esse

caso: quanto menor o tumor, menor sua agressividade. Embora se tenha observado uma coerência na evolução dos valores das características ao longo do tempo, estatisticamente não foi possível diferenciar estes valores com significância.

- **RECURR**

Por fim, para o *endpoint* RECURR, obteve-se 42 características com distribuição gaussiana e dentre elas, apenas 2 (*Contrast.1* e *Kurtosis*) apresentaram significância estatística para diferenciação das médias dos dois grupo (tabela 6.15). O gráfico de dispersão em pares indica novamente a sobreposição dos dados obtidos, o que prejudica a separação dos grupos (figura 6.11).

Tabela 6.15: Características significativas para o endpoint RECURR.

RECURR	
Características com distribuição gaussiana	42
Características com diferença significativa entre os dois grupos	p-valor
<i>Contrast.1</i>	0,016
<i>Kurtosis</i>	0,015

Figura 6.11: Gráfico de dispersão de pares para o endpoint RECURR.

Este estudo buscou uma abordagem diferenciada para estudo das características radiômicas no estudo do GBM, ela se baseou na separação das médias das características para diferenciação da SLDP e SG em escalas temporais menores, em busca de indicadores mais específicos. No caso do *endpoint* RECURR, ele apenas avalia se houve recidiva ou não no campo de tratamento, sem englobar indicação de tempo. Como na literatura ele é o endpoint mais estudado e com esta metodologia inicial não foi possível encontrar a(s) característica(s) adequada(s) para a separação dos grupos de pacientes, optamos por realizar outras duas abordagens para este conjunto de dados.

A primeira delas se baseou na redução de características significativas para a separação dos dois grupos de pacientes utilizando o agrupamento *k-means*, seguido por uma regressão logística ainda para mineração dos dados. Foram selecionadas 46

características significativas com o *k-means* e dentre elas 2 foram selecionadas na regressão logística (tabela 6.16). A primeira observação clara é que a abordagem diferenciada levou a seleção de duas características que não haviam sido selecionadas anteriormente *InterquartileRange* e *Maximum*. Com estas duas características, um modelo preditivo usando Regressão Logística com *k fold* de 5 foi desenvolvido e obteve uma AUC de 69% (figura 6.12). A matriz de confusão (figura 6.13) indica uma acurácia de 60,5%.

Tabela 6.16: Características significativas para o endpoint RECURR após a aplicação do algoritmo *k-means* e regressão logística.

RECURR	
Características selecionadas com o k-means (k=2)	46
Características selecionadas com a regressão logística	p-valor
<i>InterquartileRange</i>	0,013
<i>Maximum</i>	0,030

Figura 6.12: Curva ROC para o aprendizado de máquina utilizando a Regressão Logística como algoritmo preditor.

Figura 6.13: Matriz de Confusão para o aprendizado de máquina utilizando a Regressão Logística como algoritmo preditor.

A segunda abordagem diferenciada se baseou no trabalho de (Chaddad *et al.*, 2018), utilizando uma metodologia para redução de dimensionalidade dos dados baseada no ranking de Spearman. Com essa metodologia, 23 das 105 características iniciais mostraram-se diferentes com o índice de significância de 5%. E, com essas 23 características, aplicou-se um algoritmo de aprendizado de máquina *Random Forest* e obteve-se a seguinte curva ROC apresentada na figura 6.14.

Figura 6.141: Curva ROC para a segunda abordagem utilizando o ranking de Spearman para o endpoint *RECURR*.

Percebe-se pela figura 6.14 que a AUC do preditor é de 63%, e ao compará-lo com a literatura, aproxima-se do resultado obtido por (Ingrisch *et al.*, 2017). O estudo citado avaliou os *endpoints* como sendo SG e SLPD, com os grupos separados pela faixa temporal na mediana. Portanto, avaliamos, para nossos dados e para os *endpoints* SG e SLPD, a mesma metodologia baseada na correlação de Spearman. Para SG obtivemos 2 variáveis significativas, e para SLPD 14 características. E, com essas características, criamos um modelo preditivo utilizando RFs para comparar com

o trabalho. As acurácias estão representadas nas figuras 6.15 e 6.16, respectivamente, para SG e SLPD.

Figura 6.15: Curva ROC para a segunda abordagem utilizando o endpoint SG com a faixa temporal mediana.

Figura 6.16: Curva ROC para a segunda abordagem utilizando o endpoint SLPD com a faixa temporal mediana

Observa-se que a AUC do modelo preditivo obtida pelas figura 6.15 e 6.16 são de 56% e 48%, e quando comparado ao trabalho, o qual obteve uma AUC de 61,5% e 55% para a área necrótica do GBM, as mesmas apresentam-se na mesma ordem de grandeza. Acreditamos que a diferença entre as AUC deve-se ao fato de que, em nosso estudo, o GBM foi delineado incluindo tanto a área necrótica quanto a área de tumor ativo.

6.1 Discussões gerais

O presente estudo apresentou uma metodologia para extração de característica quantitativas das IRM usando *radiomics*. Trata-se do primeiro trabalho com uma abordagem prática simplificada, baseada no fluxo de trabalho realizado usualmente em planejamentos de radioterapia desses pacientes com tumores cerebrais, em que os volumes alvo de tratamento se baseiam no delineamento do tumor em IRM. Isso faz com que parte do fluxo de trabalho do *radiomics* seja poupado. De acordo com Narang

et al, o fluxo de Radiomics é feito por: aquisição de imagens, pré-processamento das imagens, segmentação, extração de características e análises (Narang *et al.*, 2016), sendo que com a metodologia deste estudo, as três primeiras etapas foram poupadas porque já tinham sido realizadas pelo radio oncologista. Além disso, os TPSs em radioterapia têm evoluído para permitir a interação com o usuário através de *scriptings* desenvolvidos em linguagem computacional específica, o que pode permitir o desenvolvimento de uma ferramenta direta associada ao TPS para avaliação do *radiomics* e predição de *endpoints* em tempo real para auxiliar o médico na condução do tratamento.

Este estudo buscou uma abordagem diferenciada para estudo das características radiômicas no estudo do GBM baseado na separação das médias das características para diferenciação da SLDP e SG em escalas temporais menores, em busca de indicadores mais específicos. No caso do *endpoint* RECURR, ele apenas avalia se houve recidiva ou não no campo de tratamento, sem englobar indicação de tempo.

Um ponto crucial para a seleção dos pacientes que foram incluídos neste trabalho foi a utilização da estimativa da D95% como possível preditor para os *endpoints* propostos (RECURR, SLPD e SG). Imaginamos inicialmente que a estimativa da dose recebida pelo tumor seria um dado de alto poder preditivo para a relação com o desfecho clínico. Com isso, nosso critério de inclusão foi definido para pacientes com ressecção cirúrgica parcial do volume tumoral pré-radioterapia, o que permitiria que essa avaliação fosse feita no TPS. No entanto, ao se aplicar as técnicas de redução de dimensionalidade dos dados através do teste de Shapiro-Wilk e teste-p, viu-se que essa característica não se mostrou significativa para nenhuma situação estudada. Portanto, os pacientes inicialmente excluídos (aqueles que tiveram ressecção total do tumor) poderiam ser englobados em uma nova análise, se o requisito de aquisição de IRM mudasse de pré-tratamento para pré-cirúrgico.

Este estudo iniciou-se com uma proposta de analisar estatisticamente as variáveis Radiomics para encontrar possíveis correlações entre o tratamento radioterápico e os *endpoints* clínicos,—porém, ao longo do estudo, verificou-se a possibilidade do desenvolvimento de modelos preditivos de desfecho clínico. O primeiro resultado interessante encontrado neste estudo foi a relação da *kurtosis* como uma característica

radiômica relacionada à recidiva em até 3 meses, ou seja, ao pior prognóstico estudado. Não encontramos na literatura um resultado semelhante, os trabalhos que utilizaram as características do radiomics relacionados ao GBM utilizaram abordagens distintas da utilizada neste trabalho, dois exemplos são a criação de um ranqueamento baseado nas características *radiomics* para determinação de pacientes com alto risco e baixo risco para recidiva (Ingrisch *et al.*, 2017) e a criação de normogramas para estratificação da sobrevida dos pacientes (Zhang *et al.*, 2019). Apenas encontramos uma referência sobre a kurtosis obtida pelo radiomics para auxílio na diferenciação de GBMs e linfomas cerebrais (Xiao *et al.*, 2018) e não relacionada ao prognóstico do paciente. A presença do *major axis* como característica relacionada à SLPD e à SG está de acordo com a literatura (Mazurowski *et al.*, 2014). Quando avaliamos os resultados obtidos com os modelos preditivos usando o grupo de pacientes estudado, verificamos que independente das metodologias abordadas, os resultados variam entre 60 e 70% de acurácia global dos preditores. Observamos que estes resultados são muito semelhantes aos encontrados na literatura para trabalhos analisando IRM ponderadas em T1 pós-contraste e com número de pacientes semelhantes (Ingrisch *et al.*, 2017). Metodologias mais complexas, mesclando extração de características de diversas sequências de IRM e dados clínicos foram empregadas mais recentemente e obtiveram acurácias um pouco maiores (Bae *et al.*, 2018; Chen *et al.*, 2019; Zhang *et al.*, 2019).

Por se tratar de um estudo estatístico, uma maneira para refinar os modelos desenvolvidos pode ser feita aumentando o número amostral, padronizando os parâmetros de aquisição das IRM, e até mesmo pela utilização de ferramentas automáticas para segmentação do tumor. As limitações deste estudo nestes aspectos ocorreram, pois embora a amostra utilizada tenha englobado pacientes tratados até meados de 2019, apenas pacientes em que foi possível a realização do backup dos planos de tratamento efetuados foram analisados. Alguns parâmetros das IRM utilizadas variaram entre os pacientes, embora elas tenham sido adquiridas com o protocolo padrão para pacientes com GBM do serviço de imagens do HCFMRP-USP, alguns parâmetros geométricos das imagens foram adequados pelo técnico responsável pelo exame. Já a utilização de ferramentas automáticas para segmentação do tumor foi avaliada através de um software gratuito, o BraTumIA, do inglês *Brain Tumor Image Analysis* (Rios Velazquez *et al.*, 2015), cria uma máscara binária em

quatro IRM de referência: T1, T2, FLAIR e T1 pós contraste, e a partir dessas imagens as segmentações da área necrótica, edema, e tumor ativo são feitas. No estudo citado as regiões segmentadas foram comparáveis com a segmentação automática da manual feita por radio oncologistas, mas em nossa experiência isso não ocorreu.

O presente estudo se baseou em IRM adquiridas em uma única instituição, a extrapolação dos resultados obtidos para imagens adquiridas em outras máquinas e outros locais merece cuidado, ainda há uma grande incerteza sobre qual a melhor metodologia de pré-processamento das IRM deve ser utilizado antes da aplicação do *radiomics* nelas (Um *et al.*, 2019).

7. Conclusões

Com o presente estudo foi possível extrair características Radiomics de IRM ponderadas em T1 e avaliar quais variáveis estavam mais correlacionadas com os *endpoints* SG, SLPG e RECURRE, em várias escalas temporais. Utilizando os testes t-Student, Shapiro-Wilk, *k-means* e LR para mineração de dados, observamos que as metodologias resultam em características significativas diferentes para separação dos grupos avaliados. Para SLPD em 3 meses, vimos que a *Kurtosis* pode ser utilizado como separador em classes de pacientes classificados em 0 e 1, respectivamente pacientes que tiveram recorrência da lesão dentro e fora de 3 meses. Para SG, apesar de não observarmos características significativas na maioria das escalas temporais, para 12 e 30 meses, variáveis relacionadas ao formato da lesão mostraram-se significativamente diferentes. E, para RECURRE, como é o desfecho clínico mais estudado, uma nova metodologia utilizando algoritmo de *k-means* e LR proporcionou outras duas características significativamente diferentes da metodologia baseada em teste t-Student e Shapiro-Wilk. Após a mineração dos dados, algoritmos classificadores LR e DT foram aplicados para criar modelos preditivos para os *endpoints*, e os resultados mostram-se na mesma escala de grandeza de estudos recentes. Como resultados obtivemos acurácia, para o endpoint SLPD em 3 meses: pelo modelo estatístico de 70%; acurácia utilizando LR como preditor de 88,4% (com respectiva AUC de 75%). Obteve-se AUC de 69% e acurácia global de 60,5% para o endpoint RECURRE com LR como preditor após a separação das variáveis pelo algoritmo *k-means*; e acurácia baseada na correlação de Spearman para RECURRE,

SG e SLPD, respectivamente, de 63%, 56% e 48%. Os resultados são compatíveis com a literatura. Uma maneira de refinar e aprimorar os resultados seria aumentar o número amostral, e padronizar os parâmetros de aquisição das IRM.

8. Referências bibliográficas

- ADAMSON, C. *et al.* Glioblastoma multiforme: A review of where we have been and where we are going. **Expert Opinion on Investigational Drugs**, v. 18, n. 8, p. 1061–1083, 2009.
- AERTS, H. J. W. L. *et al.* Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. **Nature Communications**, v. 5, 2014.
- AMADASUN, M.; KING, R. Textural Features Corresponding to Textural Properties. **IEEE Transactions on Systems, Man and Cybernetics**, v. 19, n. 5, p. 1264–1274, 1989.
- ASSEFA, D. *et al.* Robust texture features for response monitoring of glioblastoma multiforme on T1 -weighted and T2 -FLAIR MR images: A preliminary investigation in terms of identification and segmentation. **Medical Physics**, v. 37, n. 4, p. 1722–1736, 2010.
- AVANZO, M.; STANCANELLO, J.; NAQA, I. EL. Beyond imaging: The promise of radiomics. **Physica Medica**, v. 38, p. 122–139, 2017.
- BAE, S. *et al.* Radiomic MRI phenotyping of glioblastoma: Improving survival prediction. **Radiology**, v. 289, n. 3, p. 797–806, 2018.
- BARRINGTON, S. F. *et al.* Role of imaging in the staging and response assessment of lymphoma: Consensus of the international conference on malignant lymphomas imaging working group. **Journal of Clinical Oncology**, v. 32, n. 27, p. 3048–3058, 2014.
- BAUER, S. *et al.* Segmentation of Brain Tumor Images Based on Integrated Hierarchical Classification and Regularization. **MICCAI BraTS Workshop**, p. 10–13, 2012.
- BHOLOWALIA, P.; KUMAR, A. EBK-means: A clustering technique based on elbow method and k-means in WSN. **International Journal of Computer Applications**, v. 105, n. 9, 2014.
- BREIMAN, L.; FRIEDMAN, J. H.; OLSEN, R. A.; STONE, C. J. **Classification and Regression Trees**. [s.l.] Chapman & Hall, 1984.
- BUVAT, I.; ORLHAC, F.; SOUSSAN, M. Tumor texture analysis in PET: Where do we stand? **Journal of Nuclear Medicine**, v. 56, n. 11, p. 1642–1644, 2015.
- CABRERA, A. R. *et al.* Radiation therapy for glioblastoma: Executive summary of an American Society for Radiation Oncology Evidence-Based Clinical Practice Guideline.

- Practical Radiation Oncology**, v. 6, n. 4, p. 217–225, 2016.
- CARUANA, R.; NICULESCU-MIZIL, A. An empirical comparison of supervised learning algorithms. **ACM International Conference Proceeding Series**, v. 148, p. 161–168, 2006.
- CHADDAD, A. *et al.* Prediction of survival with multi-scale radiomic analysis in glioblastoma patients. **Medical and Biological Engineering and Computing**, v. 56, n. 12, p. 2287–2300, 2018.
- _____. Radiomics in glioblastoma: Current status and challenges facing clinical implementation. **Frontiers in Oncology**, v. 9, n. MAY, p. 1–9, 2019.
- CHEN, X. *et al.* Development and Validation of a MRI-Based Radiomics Prognostic Classifier in Patients with Primary Glioblastoma Multiforme. **Academic Radiology**, v. 26, n. 10, p. 1292–1300, 2019.
- CHICKLORE, S. *et al.* Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis. **European Journal of Nuclear Medicine and Molecular Imaging**, v. 40, n. 1, p. 133–140, 2013.
- DAVNALL, F. *et al.* Assessment of tumor heterogeneity: An emerging imaging tool for clinical practice? **Insights into Imaging**, v. 3, n. 6, p. 573–589, 2012.
- DHIMAN, D. *et al.* Collaborative Tele-Neuropsychiatry Consultation Services for Patients in Central Prisons. **Journal of Neurosciences in Rural Practice**, v. 79, n. 1, p. 68–72, 2018.
- DINSTEIN, I.; SHANMUGAM, K.; HARALICK, R. M. Textural Features for Image Classification. **IEEE Transactions on Systems, Man, and Cybernetics**, v. SMC-3, n. 6, p. 610–621, 1973.
- EISENHAUER, E. A. *et al.* New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). **European Journal of Cancer**, v. 45, n. 2, p. 228–247, 2009.
- ELIAT, P. A. *et al.* Can dynamic contrast-enhanced magnetic resonance imaging combined with texture analysis differentiate malignant glioneuronal tumors from other glioblastoma? **Neurology Research International**, v. 2012, n. i, 2012.
- GALAVIS, P. E. *et al.* Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. **Acta Oncologica**, v. 49, n. 7, p. 1012–1016, 2010.
- GIGER, M. L.; KARSSEMEIJER, N.; SCHNABEL, J. A. Breast Image Analysis for Risk Assessment, Detection, Diagnosis, and Treatment of Cancer. **Annual Review of**

Biomedical Engineering, v. 15, n. 1, p. 327–357, 2013.

GILLIES, R. J. *et al.* The biology underlying molecular imaging in oncology: from genome to anatome and back again. **Clinical Radiology**, v. 65, n. 7, p. 517–521, 2010.

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, 2016.

GRUS, J. **Data Science from Scratch**. Sebastopol,: O'Reilly, 2015.

HANIF, F. *et al.* Glioblastoma multiforme: A review of its epidemiology and pathogenesis through clinical presentation and treatment. **Asian Pacific Journal of Cancer Prevention**, v. 18, n. 1, p. 3–9, 2017.

HOU, L. C. *et al.* Recurrent glioblastoma multiforme: a review of natural history and management options. **Neurosurgical focus**, v. 20, n. 4, 2006.

HUMPHRIES, H. N. *et al.* Characterization of Cancer Stem Cells in Colon Adenocarcinoma Metastasis to the Liver. **Frontiers in Surgery**, v. 4, n. January, p. 1–10, 2018.

INGRISCH, M. *et al.* Radiomic analysis reveals prognostic information in T1-weighted baseline magnetic resonance imaging in patients with glioblastoma. **Investigative Radiology**, v. 52, n. 6, p. 360–366, 2017.

JOVČEVSKA, I.; KOČEVAR, N.; KOMEL, R. Glioma and glioblastoma - how much do we (not) know? **Molecular and Clinical Oncology**, v. 1, n. 6, p. 935–941, 2013.

KICKINGEREDER, P. *et al.* Radiomic profiling of glioblastoma: Identifying an imaging predictor of patient survival with improved performance over established clinical and radiologic risk models. **Radiology**, v. 280, n. 3, p. 880–889, 2016.

KUMAR, V. *et al.* Radiomics: The process and the challenges. **Magnetic Resonance Imaging**, v. 30, n. 9, p. 1234–1248, 2012.

LAMBIN, P. *et al.* Radiomics: Extracting more information from medical images using advanced feature analysis. **European Journal of Cancer**, v. 48, n. 4, p. 441–446, 2012.

_____. Predicting outcomes in radiation oncology-multifactorial decision support systems. **Nature Reviews Clinical Oncology**, v. 10, n. 1, p. 27–40, 2013.

LAMBIN, P. Radiomics Digital Phantom. **CancerData**, v. 41, n. 2, p. 366–373, 2016.

LOPES, I. C. Giordano Bruno: entre o geocentrismo e o heliocentrismo. **Griot : Revista de Filosofia**, v. 9, n. 1 SE-Artigos, 15 jun. 2014.

LOUIS, D. N. *et al.* The 2007 WHO classification of tumours of the central nervous system. **Acta Neuropathologica**, v. 114, n. 2, p. 97–109, 2007.

LUDWIG, J. A.; WEINSTEIN, J. N. Biomarkers in cancer staging, prognosis and

treatment selection. **Nature Reviews Cancer**, v. 5, n. 11, p. 845–856, 2005.

MACQUEEN, J.; OTHERS. **Some methods for classification and analysis of multivariate observations** Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. **Anais...**1967

MAHER, E. A. *et al.* Malignant glioma: Genetics and biology of a grave matter. **Genes and Development**, v. 15, n. 11, p. 1311–1333, 2001.

MAZUROWSKI, M. A. *et al.* Computer-extracted MR imaging features are associated with survival in glioblastoma patients. **Journal of Neuro-Oncology**, v. 120, n. 3, p. 483–488, dez. 2014.

MINNITI, G. *et al.* Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma in elderly patients. **Journal of Neuro-Oncology**, v. 88, n. 1, p. 97–103, 2008.

MOLINA-GARCÍA, D. *et al.* Prognostic models based on imaging findings in glioblastoma: Human versus Machine. **Scientific Reports**, v. 9, n. 1, p. 1–9, 2019.

NARANG, S. *et al.* Radiomics in glioblastoma: Current status, challenges and potential opportunities. **Translational Cancer Research**, v. 5, n. 4, p. 383–397, 2016.

NAVONE, S. E. *et al.* Significance and Prognostic Value of The Coagulation Profile in Patients with Glioblastoma: Implications for Personalized Therapy. **World Neurosurgery**, v. 121, p. e621–e629, 2019.

O’CONNOR, J. P. B. Cancer heterogeneity and imaging. **Seminars in Cell and Developmental Biology**, v. 64, p. 48–57, 2017.

RIOS VELAZQUEZ, E. *et al.* Fully automatic GBM segmentation in the TCGA-GBM dataset: Prognosis and correlation with VASARI features. **Scientific Reports**, v. 5, n. December 2014, p. 1–10, 2015.

SCALCO, E.; RIZZO, G. Texture analysis of medical images for radiotherapy applications. **British Journal of Radiology**, v. 90, n. 1070, 2017.

SCHAIJK, B. VAN *et al.* Circulating tumor stem cells and glioblastoma: A review. **Journal of Clinical Neuroscience**, v. 61, p. 5–9, 2019.

SHUKLA, G. *et al.* Advanced magnetic resonance imaging in glioblastoma: A review. **Chinese Clinical Oncology**, v. 6, n. 4, 2017.

SILVA, J. P. B. C. DA. Modelos de Regressão Linear e Logística utilizando o software R. p. 131, 2016.

STEINFELD, A. D. **Principles and Practice of Radiation Oncology**. 2nd ed . [s.l.: s.n.]. v. 184

STUPP, R. *et al.* Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. **The Lancet Oncology**, v. 10, n. 5, p. 459–466, 2009.

TANG, J. *et al.* Computer-Aided Detection and Diagnosis of Breast Cancer With Mammography: Recent Advances Stereotactic Body Radiation Therapy View project Radiation and Quantitative Imaging View project Computer-Aided Detection and Diagnosis of Breast Cancer With Mammogr. **Ieee Transactions on Information Technology in Biomedicine**, v. 13, n. 2, p. 236–251, 2009.

UM, H. *et al.* Impact of image preprocessing on the scanner dependence of multi-parametric MRI radiomic features and covariate shift in multi-institutional glioblastoma datasets. **Physics in Medicine and Biology**, v. 64, n. 16, 2019.

VANDENBERGHE, M. E. *et al.* Relevance of deep learning to facilitate the diagnosis of HER2 status in breast cancer. **Scientific Reports**, v. 7, 2017.

WALKER, M. D.; STRIKE, T. A.; SHELINE, G. E. An analysis of dose-effect relationship in the radiotherapy of malignant gliomas. **International Journal of Radiation Oncology, Biology, Physics**, v. 5, n. 10, p. 1725–1731, 1979.

WEHRENS, R.; KRUISSELBRINK, J. Flexible self-organizing maps in kohonen 3.0. **Journal of Statistical Software**, v. 87, n. 7, 2018.

XIAO, D.-D. *et al.* Glioblastoma and primary central nervous system lymphoma: Preoperative differentiation by using MRI-based 3D texture analysis. **Clinical Neurology and Neurosurgery**, v. 173, p. 84–90, out. 2018.

YANG, D. *et al.* Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma. **Medical Physics**, v. 42, n. 11, p. 6725–6735, 2015.

YIP, S. S. F.; AERTS, H. J. W. L. Applications and limitations of radiomics. **Physics in Medicine and Biology**, v. 61, n. 13, p. R150–R166, 2016.

ZACHARAKI, E. I. *et al.* Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. **Magnetic Resonance in Medicine**, v. 62, n. 6, p. 1609–1618, 2009.

ZAR, J. H. Significance testing of the spearman rank correlation coefficient. **Journal of the American Statistical Association**, v. 67, n. 339, p. 578–580, 1972.

ZHANG, X. *et al.* A radiomics nomogram based on multiparametric MRI might stratify glioblastoma patients according to survival. **European Radiology**, v. 29, n. 10, p.

5528–5538, out. 2019.

ZHOU, M. *et al.* Radiologically defined ecological dynamics and clinical outcomes in glioblastoma multiforme: Preliminary results. **Translational Oncology**, v. 7, n. 1, p. 5–13, 2014.