



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENERGIA NUCLEAR
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIAS ENERGÉTICAS E
NUCLEARES

DANILLO MENEZES OLIVEIRA

**HARMONIZAÇÃO DOS ATRIBUTOS RADIÔMICOS EM TOMOGRAFIA
COMPUTADORIZADA DE COVID-19**

Recife

2022

DANILLO MENEZES OLIVEIRA

**HARMONIZAÇÃO DOS ATRIBUTOS RADIÔMICOS EM TOMOGRAFIA
COMPUTADORIZADA DE COVID-19**

Dissertação apresentada ao Programa de Pós-Graduação em Tecnologias Energéticas e Nucleares da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para obtenção do título de Mestre em Tecnologias Energéticas e Nucleares. Área de concentração: Dosimetria e Instrumentação Nuclear.

Orientador: Fernando Roberto de Andrade Lima

Coorientadora: Mércia Liane de Oliveira

Recife

2022

Catálogo na fonte:
Bibliotecário Carlos Moura, CRB-4 / 1502

- O48h Oliveira, Danilo Menezes.
Harmonização dos atributos radiômicos em tomografia computadorizada de COVID-19. / Danilo Menezes Oliveira. – 2022.
88 f.: il.
- Orientador: Prof. Dr. Fernando Roberto de Andrade Lima.
Coorientadora: Profa. Dra. Mércia Liane de Oliveira.
Dissertação (mestrado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Tecnologias Energéticas e Nucleares, 2022.
Inclui referências, apêndices e anexo.
1. Tomografia computadorizada. 2. COVID-19. 3. Radiômica.
4. Harmonização. I. Lima, Fernando Roberto de Andrade (orientador).
II. Oliveira, Mércia Liane de (coorientadora). III. Título.

616.07572 (22. ed.)

UFPE
BCTG/2022-191

DANILLO MENEZES OLIVEIRA

**HARMONIZAÇÃO DOS ATRIBUTOS RADIÔMICOS EM TOMOGRAFIA
COMPUTADORIZADA DE COVID-19**

Dissertação apresentada ao Programa de Pós-Graduação em Tecnologias Energéticas e Nucleares da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para obtenção do título de Mestre em Tecnologias Energéticas e Nucleares. Área de concentração: Dosimetria e Instrumentação Nuclear.

Aprovado em: 23/02/2022.

BANCA EXAMINADORA

Prof. Dr. Vinícius de Oliveira Menezes (Examinador Externo)
Universidade Federal de Pernambuco - UFPE

Prof^a. Dr^a. Ana Figueiredo Maia (Examinadora Externa)
Universidade Federal de Sergipe - UFS

Prof^a. Dr^a. Viviane Khoury Asfora (Examinadora Interna)
Universidade Federal de Pernambuco - UFPE

AGRADECIMENTOS

Primeiramente agradeço a Deus, pela saúde e pela oportunidade da vida. Aos meus pais, por me ensinarem os princípios éticos e morais, e por me apoiar em todos os momentos da minha vida. Agradeço aos colaboradores do Hospital Universitário Alcides Carneiro (HUAC/UFCG), em especial a minha chefe Ma. Ana Maria Cabral, pelo apoio na execução deste trabalho durante o período em que trabalhei neste hospital. Aos pesquisadores do Hospital Universitário de Juiz de Fora (HUJF/UFJF), do Hospital Universitário Alcides Carneiro (HUAC/UFCG) e do Complexo Hospitalar Universitário Prof. Edgard Santos (HUPES/UFBA). Aos integrantes do nosso grupo de pesquisa, em especial ao Me. Marcos Machado, pelo incentivo a sempre ir além, pautados com produtivas discursões, e ao Me. Mauro Namías, pelo apoio e ensinamentos relativos as harmonizações. Aos professores e pesquisadores do Departamento de Energia Nuclear (DEN/UFPE) e do Centro Regional de Ciências Nucleares do Nordeste (CRCN-NE/CNEN), em especial ao Prof. Dr. Fernando de Andrade Lima e a Profa. Dra. Mércia Oliveira pela orientação e auxílio na construção deste trabalho. Agradeço a Financiadora de Estudos e Pesquisas (Finep) pelo financiamento do trabalho desenvolvido pelo grupo de pesquisa.

RESUMO

Com o surgimento da COVID-19 em 2019, o grande desafio foi o diagnóstico inicial da doença como forma de atenuar a propagação do vírus SARS-CoV-2. Exames de tomografia computadorizada (TC) de tórax podem ser utilizados tanto para o diagnóstico quanto para o estadiamento da COVID-19. Entretanto, o seu diagnóstico pode ser confundido com outras infecções agudas. Para melhorar a acurácia diagnóstica, sistemas computadorizados utilizando inteligência artificial (IA) combinado com radiômica podem ser desenvolvidos. Este trabalho integra um projeto de pesquisa cujo objetivo principal é o desenvolvimento e validação de um software para diagnóstico de COVID-19 baseado em TC, radiômica e IA, no qual participam três centros de radiologia. Entretanto, na condução de estudos radiômicos multicêntricos, as diferenças de modelos de equipamentos, protocolos de aquisição e de processamento das imagens podem influenciar os atributos radiômicos obtidos. Uma maneira de reduzir essa variabilidade é a harmonização dos atributos radiômicos entre os centros de pesquisa. Para este trabalho, os atributos dos três centros de radiologia foram selecionados por métodos supervisionado e não supervisionado. Adicionalmente, foi realizada a avaliação da correlação do atributo com o volume das lesões. Aqueles atributos pouco correlacionados ao volume foram harmonizados utilizando o método ComBat. Observou-se que o método não supervisionado, além de selecionar atributos correlacionados a volume, é eficiente na triagem de atributos harmonizados, não sendo necessário uma harmonização matemática (método ComBat). Já os atributos selecionados pelo método supervisionado possuem fraca correlação com o volume da lesão, por este motivo foram harmonizados utilizando o método ComBat. Como resultado, notou-se que o método ComBat é eficiente no alinhamento dos dados dos atributos radiômicos que tem pouca correlação com o volume da lesão, tornando assim viável o estudo radiômico multicêntrico.

Palavras-chave: tomografia computadorizada; COVID-19; radiômica; harmonização.

ABSTRACT

With the emergence of COVID-19 in 2019, the great challenge was the initial diagnosis of the disease as a way to mitigate the spread of the SARS-CoV-2 virus. Computed tomography (CT) scans of the chest can be used for both diagnosis and staging of COVID-19. However, its diagnosis can be confused with other acute infections. To improve diagnostic accuracy, computerized systems using artificial intelligence (AI) combined with radiomics can be developed. This work is part of a research project whose main objective is the development and validation of a software for the diagnosis of COVID-19 based on CT, radiomics and AI, in which three radiology centers participate. However, when conducting multicenter radiomic studies, differences in equipment models, image acquisition and processing protocols can influence the features radiomics obtained. One way to reduce this variability is the harmonization of features radiomics between research centers. For this work, the features of the three radiology centers were selected by supervised and unsupervised methods. Additionally, the correlation of the features with the volume of lesions was evaluated. Those features that were weakly correlated to volume were harmonized using the ComBat method. It was observed that the unsupervised method, in addition to selecting features correlated to volume, is efficient in the screening of harmonized attributes, not requiring a mathematical harmonization (ComBat method). The features selected by the supervised method have a weak correlation with the volume of the lesion, for this reason they were harmonized using the ComBat method. As a result, it was noticed that the ComBat method is efficient in aligning the data of the features radiomics that have a weak correlation with the lesion volume, thus making the multicentric radiomic study viable.

Keywords: computed tomography; COVID-19; radiomics; harmonization.

LISTA DE FIGURAS

Figura 1 –	Imagem tomográfica reconstruída no plano axial, coronal e sagital	16
Figura 2 –	Ilustração do princípio de formação da imagem tomográfica	17
Figura 3 –	Cada célula numa matriz de imagem de uma TC é uma representação bidimensional (pixel) de um volume de tecido (voxel)	18
Figura 4 –	Ilustração dos passos de uma aquisição tomográfica axial	21
Figura 5 –	Ilustração dos passos de uma aquisição tomográfica helicoidal	21
Figura 6 –	A matriz 3 x 3 no centro da figura representa os valores reais de atenuação e os valores das projeções (a, b, c, d) mostram o somatório dos valores de atenuação para este exemplo	23
Figura 7 –	Ilustração da sequência da reconstrução iterativa de imagens tomográficas	24
Figura 8 –	Tomografia de tórax de paciente do sexo feminino, de 34 anos, com dor torácica e tosse há 4 dias e RT-PCR positivo para SARS-CoV-2. A imagem apresenta opacidades pulmonares com atenuação de vidro fosco, de predomínio periférico, bilaterais	26
Figura 9 –	Tomografia de tórax de paciente do sexo masculino, de 52 anos, com anosmia e tosse seca há 6 dias e RT-PCR positivo para SARS-CoV-2. A imagem apresenta padrão pulmonar de pavimentação em mosaico bilateral, com opacidades em vidro fosco associadas a espessamentos septais interlobulares	26
Figura 10 –	Tomografia de tórax de paciente do sexo masculino, de 63 anos, com dispneia e episódios de febre há 7 dias, com RT-PCR positivo para SARS-CoV-2. A imagem apresenta consolidações pulmonares (setas) periféricas bilaterais	27
Figura 11 –	Tomografia de tórax de paciente do sexo feminino, de 71 anos, com dispneia e tosse seca há 8 dias, com RT-PCR positivo para SARS-CoV-2. A imagem apresenta opacidades reticulares (setas) pulmonares basais periféricas bilaterais	27
Figura 12 –	Tomografia de tórax mostrando linha subpleural (setas) em paciente com diagnóstico confirmado de COVID-19, com tosse seca, dispneia e anosmia há 9 dias	28
Figura 13 –	Tomografia de tórax de paciente do sexo masculino, de 59 anos, internado na unidade de terapia intensiva há 12 dias com diagnóstico de COVID-19. A imagem apresenta opacidades	28

pulmonares, uma delas com sinal do halo invertido (setas), caracterizado por consolidação com centro de baixa atenuação

Figura 14 – Tomografia de tórax mostrando sinal do halo caracterizado por consolidação nodular pulmonar com halo de vidro fosco (setas) em paciente com diagnóstico de COVID-19	29
Figura 15 – Fluxograma dos processos radiômicos	38
Figura 16 – Layout da interface do Lifex ilustrando três segmentações de volumes de interesse em uma tomografia de tórax	40
Figura 17 – Gráfico da distribuição normal dos volumes das lesões em voxels	43
Figura 18 – Gráficos de dispersões dos atributos normalizados do a. Atributo 01, b. Atributo 02, c. Atributo 03, d. Atributo 04, e. Atributo 05, f. Atributo 06, g. Atributo 07, h. Atributo 08 e i. Atributo 09 em função do volume da lesão normalizado	47
Figura 19 – Gráfico da distribuição normal do Atributo 01	49
Figura 20 – Gráfico de distribuição normal do Atributo 02	50
Figura 21 – Gráfico de distribuição normal do Atributo 03	51
Figura 22 – Gráfico de distribuição normal do Atributo 04	53
Figura 23 – Gráfico de distribuição normal do Atributo 05	54
Figura 24 – Gráficos de distribuição normal do Atributo 06	56
Figura 25 – Gráficos de distribuição normal do Atributo 07	57
Figura 26 – Gráfico de distribuição normal do Atributo 08	59
Figura 27 – Gráfico de distribuição normal do Atributo 09	60

LISTA DE TABELAS

Tabela 1 – Características dos equipamentos	39
Tabela 2 – Características dos exames obtidos	39
Tabela 3 – Atributos selecionados	45
Tabela 4 – Dados do Atributo 01	49
Tabela 5 – Dados do Atributo 02	50
Tabela 6 – Dados do Atributo 03	51
Tabela 7 – Dados do Atributo 04	52
Tabela 8 – Dados do Atributo 05	54
Tabela 9 – Dados do Atributo 06	55
Tabela 10 – Dados do Atributo 07	57
Tabela 11 – Dados do Atributo 08	58
Tabela 12 – Dados do Atributo 09	60

LISTA DE ABREVIATURAS E SIGLAS

Bit	<i>Binary Digit</i>
COVID-19	<i>Coronavirus Disease 2019</i>
CSV	<i>Comma separated values</i>
DNA	Ácido desoxirribonucleico
FOV	<i>Field of view</i>
GLCM	<i>Gray-Level Cooccurrence Matrix</i>
GLDM	<i>Gray-Level Distance Matrix</i>
GLDZM	<i>Gray-Level Distance Zone Matrix</i>
GLRLM	<i>Gray-Level Run-length Matrix</i>
GLSZM	<i>Gray-Level Size Zone Matrix</i>
H1N1	Influenza A subtipo H1N1
HU	<i>Hounsfield unit</i>
HUAC	Hospital Universitário Alcides Carneiro
HUJF	Hospital Universitário de Juiz de Fora
HUPES	Hospital Universitário Professor Edgar Santos
IA	Inteligência artificial
ICC	<i>Intraclass correlation coefficient</i>
IgG	Imunoglobulina G
IgM	Imunoglobulina M
LAPI	Laboratório de Pesquisa em Infectologia
MERS	Síndrome respiratória do Médio Oriente
NGLDM	<i>Neighborhood Gray-Level Dependence Matrix</i>
NGTDM	<i>Neighborhood Gray-Tone Difference Matrix</i>
OMS	Organização Mundial da Saúde
PACS	<i>Picture Archiving and Communication System</i>
PET	<i>Positron Emission Tomography</i>
RNA	Ácido ribonucleico
ROC	Receiver Operating Characteristic
ROI	<i>Region of interest</i>
RT-PCR	<i>Reverse transcription polymerase chain reaction</i>
SARS	<i>Severe Acute Respiratory Syndrome</i>
SARS-CoV-2	<i>Severe Acute Respiratory Syndrome Coronavirus 2</i>
SRAG	Síndrome Respiratória Aguda Grave
SUV	<i>Standardized Uptake Value</i>
SVA	<i>Surrogate Variable Analysis</i>
TC	Tomografia Computadorizada
UFBA	Universidade Federal da Bahia
UFCG	Universidade Federal de Campina Grande
UFJF	Universidade Federal de Juiz de Fora
VOI	<i>Volume of interest</i>

SUMÁRIO

1	INTRODUÇÃO	13
2	REVISÃO DE LITERATURA	16
2.1	TOMOGRAFIA COMPUTADORIZADA (TC)	16
2.1.1	Princípios básicos de tomografia computadorizada	17
2.1.1.1	Matriz da imagem tomográfica	18
2.1.1.2	Unidade de hounsfield (HU)	19
2.1.2	Aquisição das imagens tomográficas	20
2.1.2.1	Aquisição axial	20
2.1.2.2	Aquisição helicoidal	21
2.1.3	Processamento das imagens tomográficas	22
2.1.3.1	Retroprojeção filtrada	22
2.1.3.2	Reconstrução iterativa	23
2.2	DIAGNÓSTICO DA COVID-19 POR TOMOGRAFIA COMPUTADORIZADA	24
2.2.1	Principais achados radiológicos dos exames de tomografia computadorizada no diagnóstico da COVID-19	25
2.2.2	Desafios no diagnóstico da COVID-19 por tomografia computadorizada	29
2.3	RADIÔMICA	30
2.3.1	Atributos radiômicos	31
2.3.1.1	Atributos baseados em histogramas	31
2.3.1.2	Atributos baseados em textura	31
2.3.1.3	Atributos baseados em modelos	33
2.3.1.4	Atributos baseados em formas	34
2.3.1.5	Atributos baseados em transformadas	34
2.3.1.6	Seleção (qualificação) dos atributos radiômicos	34
2.3.2	Harmonização dos atributos radiômicos	35
3	METODOLOGIA	38
3.1	AQUISIÇÃO DAS IMAGENS	38
3.2	SEGMENTAÇÃO DAS REGIÕES DE INTERESSE E EXTRAÇÃO DOS ATRIBUTOS RADIÔMICOS	40

3.3	SELEÇÃO E HARMONIZAÇÃO DOS ATRIBUTOS RADIÔMICOS	41
3.3.1	Método não supervisionado de seleção	41
3.3.2	Método supervisionado de seleção	42
3.3.3	Avaliação da correlação do atributo com o volume da lesão	42
3.3.4	Método de harmonização ComBat	43
4	RESULTADOS E DISCUSSÃO	45
4.1	SELEÇÃO DOS ATRIBUTOS RADIÔMICOS	45
4.2	HARMONIZAÇÃO DOS ATRIBUTOS RADIÔMICOS	48
4.2.1	Análise dos atributos selecionados pelo método não supervisionado	48
4.2.2	Análise dos atributos selecionados pelo método supervisionado	52
5	CONCLUSÃO	63
	REFERÊNCIAS	64
	APÊNDICE A – CÓDIGO PYTHON DE HARMONIZAÇÃO DOS ATRIBUTOS RADIÔMICOS	66
	APÊNDICE B – LISTA DOS ATRIBUTOS RADIÔMICOS	68
	ANEXO A – PARECER CONSUBSTANCIADO DO CEP	71

1 INTRODUÇÃO

Em dezembro de 2019, diversos casos de pneumonia por causa desconhecida surgiram na cidade de Wuhan, província de Hubei, China. A partir da análise do material genético isolado do vírus, constatou-se que se tratava de um novo betacoronavírus, inicialmente denominado 2019-nCoV pela Organização Mundial da Saúde (OMS). Em seguida, esse passou a ser chamado de SARS-CoV-2 (do inglês *Severe Acute Respiratory Syndrome Coronavirus 2*). A infecção pelo vírus SARS-CoV-2 causa a COVID-19 (do inglês, *Coronavirus Disease 2019*), cujos principais sintomas são febre, fadiga e tosse seca, podendo evoluir para dispneia ou, em casos mais graves, Síndrome Respiratória Aguda Grave (SRAG). A doença espalhou-se rapidamente pelo território chinês e, posteriormente, pacientes infectados por SARS-CoV-2 foram identificados em todo o mundo.

Em 30 de janeiro de 2020, a OMS declarou a doença como uma emergência de saúde pública global e, em 11 de março de 2020, ela passou a ser considerada uma pandemia (BRASIL, 2020). No mundo, até 10 de outubro de 2021, a Organização Mundial da Saúde (OMS) havia relatado mais de 237,4 milhões de casos confirmados e mais de 4,8 milhões de óbitos por COVID-19. No Brasil, até 15 de outubro de 2021, 21.612.237 casos e 602.099 óbitos por COVID-19 foram confirmados, sendo estimada uma taxa de incidência acumulada de 10.284,3 casos por 100 mil habitantes, e uma taxa de mortalidade acumulada de 286,5 óbitos por 100 mil habitantes (BRASIL, 2021).

O diagnóstico da COVID-19 pode ser realizado por testes de diagnóstico laboratorial como a reação em cadeia da polimerase por transcriptase reversa (RT-PCR), que amplifica sequências de RNA do vírus, possibilitando sua identificação. Outros testes também recomendados para o diagnóstico são os testes sorológicos com identificação de anticorpos IgM e IgG ao SARS-CoV-2, mas não são recomendados para a confirmação diagnóstica de pacientes com sintomas de início recente (BRASIL, 2020). Além disso, o diagnóstico também pode ser realizado pela detecção do antígeno do SARS-CoV-2, onde a sua principal vantagem é o resultado imediato, entretanto, sem fins de diagnóstico confirmatório (SANTOS et al., 2021).

Um método de imagem que é recomendado para o diagnóstico da COVID-19 é o exame de Tomografia Computadorizada (TC) de tórax, que também tem se mostrado útil no prognóstico de pacientes. Entretanto, nesses exames de pacientes infectados

com o SARS-CoV-2 o diagnóstico pode ser confundido com o de outras infecções agudas como influenza, SARS, MERS e H1N1 (ZU et al., 2020). Para auxiliar médicos radiologistas no diagnóstico da COVID-19 por meio de exames de TC de tórax, sistemas computadorizados podem ser desenvolvidos utilizando inteligência artificial (IA) combinado com radiômica (do inglês, *radiomics*) com o objetivo aumentar a acurácia diagnóstica.

A radiômica, como uma ferramenta analítica quantitativa, pode quantificar fenótipos de imagem, extraindo características específicas de imagens médicas. Uma assinatura radiômica de TC de tórax pode ser eficaz no diagnóstico e na predição prognóstica da COVID-19 (WU et al., 2020). Um estudo chinês (FANG et al., 2020) demonstrou potencial para a aplicação da radiômica com inteligência artificial como método para diagnóstico da COVID-19.

O número de publicações envolvendo análises radiômicas com estudos de imagens médicas tem crescido rápido. Entretanto, na condução de estudos multicêntricos, as diferenças de modelos de equipamentos, protocolos de aquisição e de processamento das imagens podem influenciar os atributos radiômicos obtidos. Essa variabilidade em estudos radiômicos implica que o modelo radiômico estabelecido usando dados de um centro pode não ser diretamente aplicável aos dados de outra instituição. Essa é uma grande limitação na disseminação dos modelos radiômicos e de sua transferência para a prática clínica (ORLHAC et al., 2018).

Uma maneira de reduzir a variabilidade dos valores dos atributos radiômicos intrínsecos aos estudos multicêntricos é a harmonização dos dados entre os centros de pesquisa. No estudo realizado por Orhac e colaboradores (2018) foi demonstrado que o método de harmonização ComBat, que é um método inicialmente utilizado para análises genômicas, conseguiu remover as variações dos valores dos atributos radiômicos e das medidas de SUV (*Standardized Uptake Value*) derivadas de imagens de PET (*Positron Emission Tomography*) adquiridas em diferentes centros e sob diferentes condições.

A harmonização dos dados dos atributos radiômicos é necessária para que seja garantida a eficiência e a aplicabilidade clínica dos estudos radiômicos multicêntricos utilizando imagens de TC na detecção de COVID-19. O objetivo do presente trabalho é harmonizar os dados dos atributos radiômicos obtidos em um estudo multicêntrico

de imagens de TC de tórax de pacientes com pneumonia. Os objetivos específicos são:

- Selecionar e extrair os dados dos atributos radiômicos de exames de TC de tórax de três centros de radiologia diagnóstica;
- Quando aplicável, harmonizar os atributos radiômicos utilizando o método ComBat;
- Avaliar a harmonização dos atributos radiômicos selecionados.

Os dados otimizados dos atributos serão utilizados no desenvolvimento de um sistema computadorizado para o diagnóstico de COVID-19 baseado em TC, radiômica e inteligência artificial.

2 REVISÃO DE LITERATURA

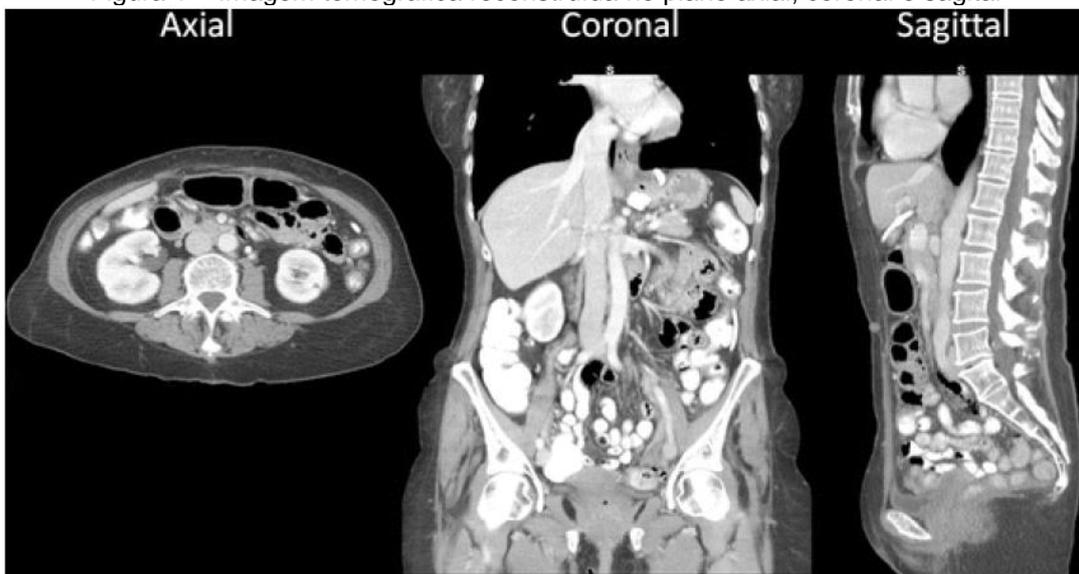
Neste capítulo, são abordados os princípios de funcionamento tecnológico da TC e suas características de desempenho. Em seguida, é apresentado o uso clínico da TC no diagnóstico da COVID-19. Por último, são apresentadas as principais características da radiômica, além de abordar suas dificuldades em estudos multicêntricos.

2.1 TOMOGRAFIA COMPUTADORIZADA (TC)

O uso clínico de exames de TC apresentou um grande crescimento nos últimos tempos, principalmente devido aos avanços significativos na qualidade das imagens associada à redução do tempo de aquisição delas.

As imagens dos exames de TC são realizadas por meio de uma varredura ao longo do eixo z (longitudinal) do paciente, para tanto, o sistema tubo de raios X e conjunto de detectores no *gantry* giram rapidamente ao redor do paciente. Essa geometria de aquisição leva a um plano de reconstrução axial da imagem. Por este motivo, o plano axial corresponde ao plano de reconstrução principal das imagens tomográficas, sendo os planos coronais e sagitais reconstruídos a partir das imagens axiais (BUSHBERG, 2021).

Figura 1 – Imagem tomográfica reconstruída no plano axial, coronal e sagital



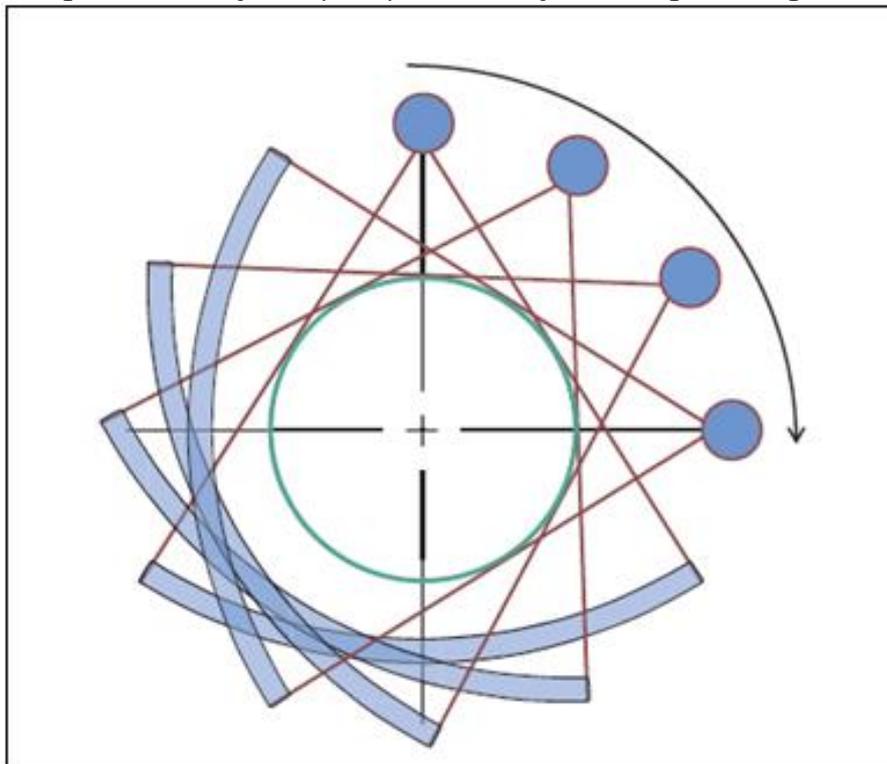
Fonte: BUSHBERG, (2021).

2.1.1 Princípios básicos de tomografia computadorizada

Quando o sistema tubo de raios X/conjunto de detectores realiza a rotação ao longo do paciente, as estruturas internas do corpo atenuam o feixe de raios X conforme suas densidades de massa e números atômicos efetivos. A intensidade da radiação detectada varia de acordo com o padrão de atenuação, e um perfil de intensidade, ou projeção, é formado.

Esse processo é repetido várias vezes e muitas projeções são produzidas, conforme ilustrado na Figura 2. Essas projeções não são apresentadas, mas armazenadas digitalmente em computadores. O processamento computacional dessas projeções envolve a superposição efetiva de cada projeção para reconstruir uma imagem das estruturas anatômicas naquele corte. O sinal do detector a cada translação tem tamanho dinâmico de 12 bits (4.096 níveis de cinza). O valor de cada incremento relaciona-se ao coeficiente de atenuação total do caminho percorrido pelos raios X através do tecido. Por meio da utilização de equações simultâneas, obtém-se uma matriz de valores, que representa a anatomia de uma secção transversal (BUSHONG, 2008).

Figura 2 - Ilustração do princípio de formação da imagem tomográfica



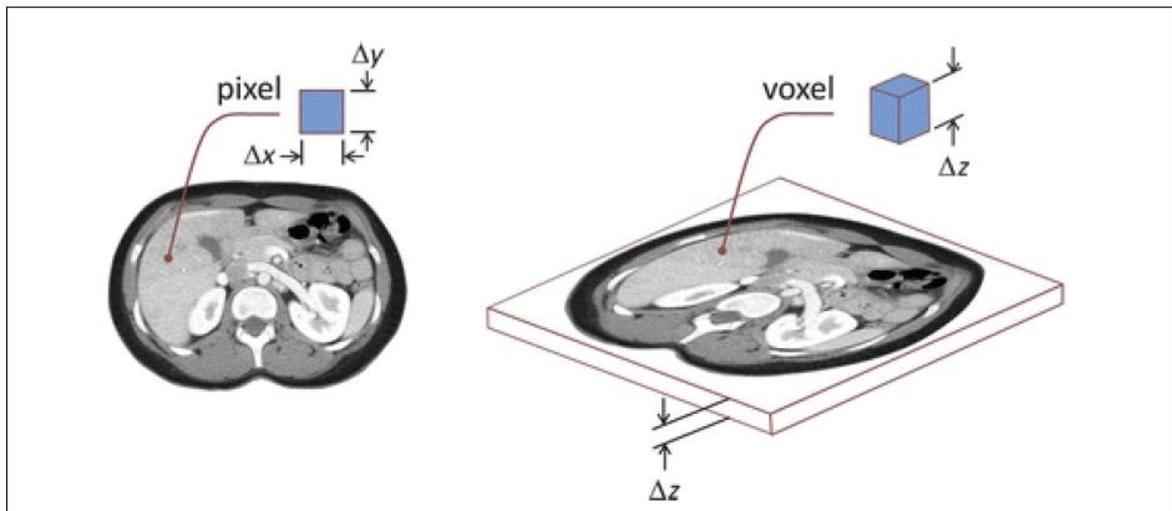
Fonte: BUSHBERG, (2012).

2.1.1.1 Matriz da imagem tomográfica

O formato da imagem tomográfica consiste em várias células, cada uma associada a um número e é apresentada com uma densidade óptica ou nível de brilho. O formato original do primeiro tomógrafo consistia em uma matriz 80×80 , com um total de 6.400 células individuais de informação. Os sistemas atuais fornecem matrizes de 512×512 , resultando em 262.144 células de informação.

Cada célula de informação é um pixel (do inglês, *picture element*) e a informação numérica contida em cada pixel é um número de TC, ou unidade hounsfield (HU, do inglês, *hounsfield unit*). O pixel é uma representação bidimensional de um volume de tecido correspondente. Já o elemento de volume do tecido é conhecido como voxel (do inglês, *volume element*); ele é determinado pelo produto entre o tamanho do pixel e a espessura do corte da imagem tomográfica (Figura 3) (BUSHONG, 2008).

Figura 3 - Cada célula numa matriz de imagem de uma TC é uma representação bidimensional (pixel) de um volume de tecido (voxel)



Fonte: BUSHBERG, (2021).

O diâmetro da reconstrução da imagem é conhecido de campo de visão (FOV, do inglês, *field of view*). Quando o FOV é aumentado por um tamanho fixo de matriz, por exemplo, de 12 cm para 20 cm, o tamanho de cada pixel é aumentado proporcionalmente. Quando se aumenta o tamanho da matriz para um FOV fixo, por exemplo, 512×512 para 1024×1024 , o tamanho do pixel diminui (BUSHONG, 2008).

Entretanto, com a diminuição do tamanho do pixel aumenta-se o ruído na imagem, por isso é necessário o balanceamento dos parâmetros imagem tomográfica.

2.1.1.2 Unidade de hounsfield (HU)

A escala de tons cinza em imagens tomográfica possui um valor quantitativo significativo. Os valores de tons de cinza em tomografia são chamados de Unidades de Hounsfield (HU), em homenagem a Godfrey Hounsfield, que foi um dos principais inovadores da tecnologia da tomografia computadorizada. O HU é definido como:

$$HU(x, y, z) = 1.000 \frac{(\mu(x, y, z) - \mu_w)}{\mu_w} \quad \text{Eq. 1}$$

onde $HU(x, y, z)$ representa a escala de tons de cinza em Unidades de Hounsfield na coordenada espacial (x, y, z) , $\mu(x, y, z)$ é o coeficiente de atenuação linear médio para um elemento de volume (voxel) de tecido do paciente na coordenada espacial (x, y, z) , e μ_w é o coeficiente de atenuação linear da água para o espectro de raios X utilizado.

Quando o coeficiente de atenuação linear médio $\mu(x, y, z)$ corresponde ao voxel da água, o numerador da Eq. 1 é zero, então o HU para a água é 0. Além disso, como $\mu(x, y, z)$ para o ar é aproximadamente zero, a razão da Eq. 1 torna-se -1, portanto, o HU para o ar é -1.000. Na prática, os HU's dos tecidos adiposos variam tipicamente de -80 a -30 e a maioria do parênquima de órgãos varia de aproximadamente +30 a +220 (HU). Os tecidos com contraste iodado e osso podem ocorrer com mais alto contraste, indo a um máximo de 3.095 HU para a maioria dos tomógrafos de 12-bit. A urina, na bexiga urinária, por ser predominantemente água, deve possuir um HU próximo de 0 e, os pulmões, por ser predominantemente ar, deve possuir um HU próximo de -1.000.

Em todas as tensões dos tubos de raios X, o HU da água é calibrado para 0 e o HU do ar para -1.000, entretanto, devido a dependência energética da atenuação, os valores de HU para tecido adiposo e tecido mole mudarão ligeiramente para diferentes tensões do tubo. Essa mudança deve ser mais significativa para materiais mais densos, como o contraste iodado, ou tecido mais densos, como osso (BUSHBERG, 2012).

2.1.2 Aquisição das imagens tomográficas

Uma vez o paciente deitado na mesa de exames, o operador do equipamento move a mesa para a abertura do *gantry* e é realizada uma varredura preliminar chamada de radiografia de localização. A depender do fabricante do equipamento de tomografia computadorizada essa radiografia inicial pode ser chamada de *scout*, topograma, escanograma ou imagem de localização. Essa imagem é utilizada para que os operadores dos equipamentos façam as marcações anatômicas inicial e final dos cortes tomográficos que serão utilizados no exame, além de determinar o tamanho do campo de visão (FOV) utilizado no estudo. Mais recentemente, alguns fabricantes de tomógrafos computadorizados utilizam essa radiografia de localização para realizar a modulação da corrente utilizada nos cortes tomográficos.

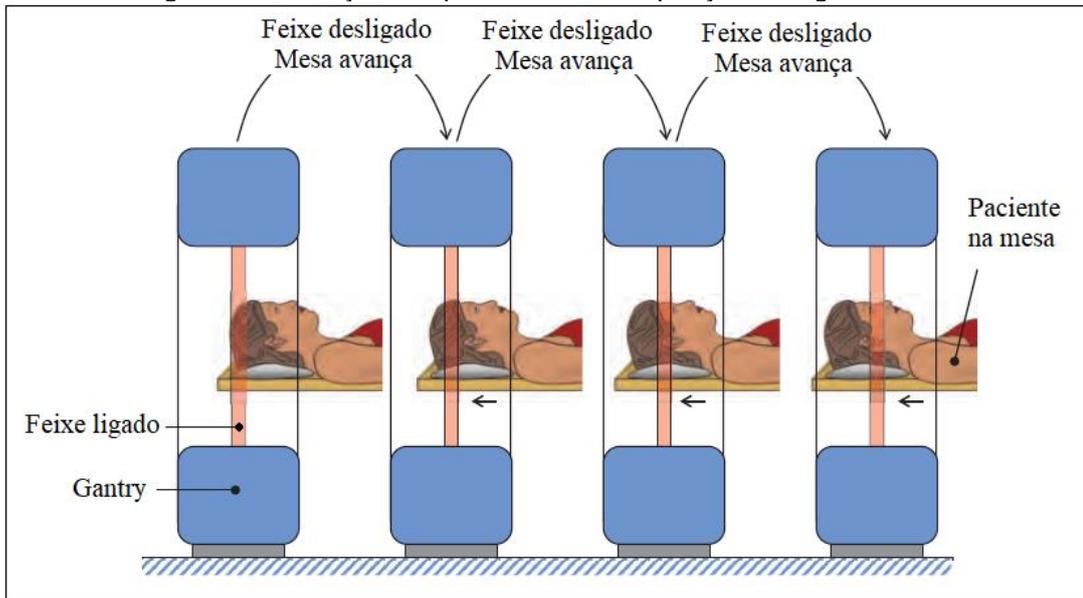
Em seguida, os parâmetros de aquisição dos cortes tomográficos são definidos, geralmente usando protocolos predefinidos. Entretanto, é necessário que o operador do equipamento faça uma análise antes das aquisições respeitando as características de cada paciente e do tipo de exame. Os parâmetros incluem tensão do feixe, corrente, tempo de rotação do *gantry*, direção da varredura, inclinação do *gantry*, configuração do detector e da colimação, parâmetros de modulação da corrente, tipos de reconstrução e o tipo de varredura que será realizada na aquisição.

Após as definições de localização dos cortes tomográficos e dos parâmetros de aquisição das imagens, a varredura tomográfica é executada. A aquisição pode ser realizada de formas distintas, como a aquisição axial (ou sequencial) e a aquisição helicoidal (ou espiral) (BUSHBERG, 2021).

2.1.2.1 Aquisição axial

A aquisição axial, também chamada de aquisição sequencial, é realizada no modo *step-and-shoot*, onde a mesa com o paciente fica parada e o *gantry* faz uma rotação de 360° para adquirir a imagem. Em seguida o feixe de radiação é desligado e mesa avança para a posição do próximo corte tomográfico. Este processo é repetido até que toda a área anatômica seja coberta, conforme ilustrado na Figura 4 (BUSHBERG, 2021).

Figura 4 – Ilustração dos passos de uma aquisição tomográfica axial

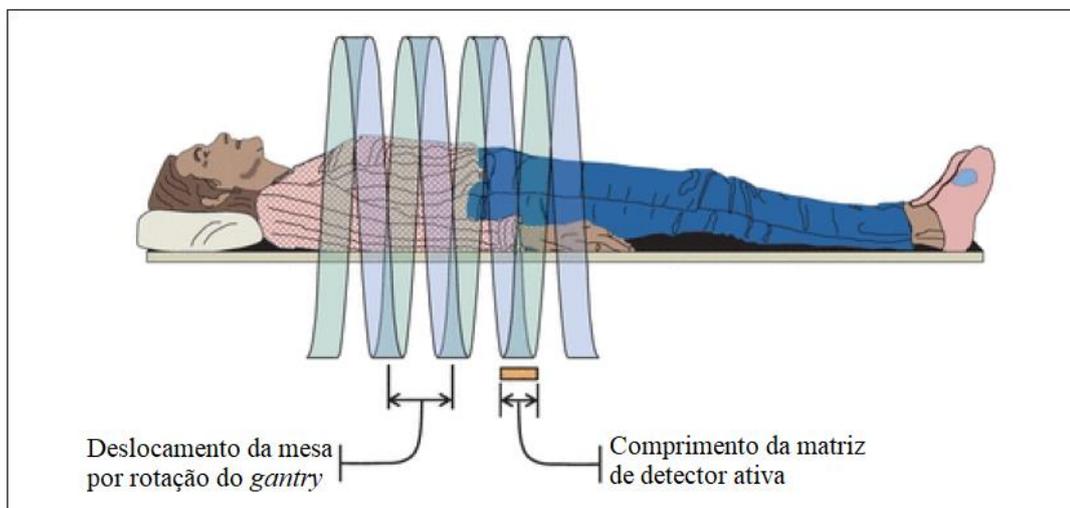


Fonte: Adaptado de BUSHBERG, (2012).

2.1.2.2 Aquisição helicoidal

Com a aquisição tomográfica helicoidal, também chamada de espiral, a mesa move-se com velocidade constante enquanto o *gantry* rotaciona em torno do paciente. Esta geometria resulta na fonte de raios X formando uma hélice ao redor do paciente, conforme ilustrado da Figura 5. A vantagem da aquisição helicoidal é a velocidade, pois não é necessário o desligamento do feixe para a mover a mesa, conforme ocorre nas aquisições axiais (BUSHBERG, 2012).

Figura 5 – Ilustração dos passos de uma aquisição tomográfica helicoidal



Fonte: Adaptado de BUSHBERG, (2021).

Semelhantemente a um cabo espiral de um telefone fixo, o *pitch* descreve o avanço relativo da mesa por rotação do *gantry*. O *pitch* de varredura helicoidal é definido como:

$$Pitch = \frac{F_{table}}{nT} \quad \text{Eq. 2}$$

onde F_{table} é o deslocamento da mesa por rotação do gantry e nT é a espessura nominal do feixe colimado. Em princípio, um *pitch* de 1,0 corresponde a uma aquisição tomográfica axial contínua; um *pitch* inferior a 1,0 resulta em uma sobre-exposição de área irradiada, conseqüentemente uma maior exposição à radiação para o paciente, enquanto um *pitch* superior a 1,0 resulta em uma subexposição da área irradiada, conseqüentemente uma exposição à radiação menor para o paciente.

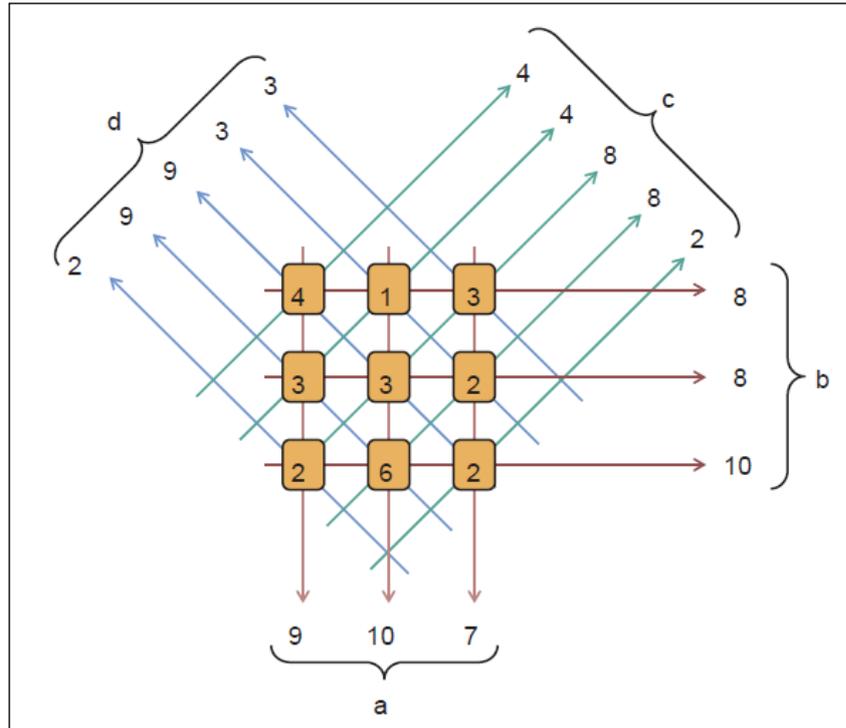
2.1.3 Processamento das imagens tomográficas

Com as projeções adquiridas, é necessário processá-las computacionalmente para serem assim apresentadas em cortes seccionais dos pacientes. Os principais tipos de processamento de imagens tomográficas utilizados atualmente são a retroprojeção filtrada e reconstrução iterativa.

2.1.3.1 Retroprojeção filtrada

As projeções adquiridas podem ser retroprojetadas com o intuito de obter os dados de coeficiente de atenuação linear em cada pixel da matriz da imagem. Na Figura 6, está ilustrado o princípio básico da retroprojeção, considerando-se uma matriz 3 x 3 representando um objeto de imagem, onde quatro projeções através desse “paciente” são apresentadas (a-d). É observado que o valor projetado ao longo de cada linha que passa pelo paciente é a soma linear dos valores de voxel naquela linha. Para uma matriz de 3 x 3, o cálculo retroprojetado é simples, entretanto uma matriz típica de uma imagem tomográfica é de 512 x 512 e a tarefa de retroprojetar os valores na imagem a partir dos valores de projeção é mais complexa (BUSHBERG, 2021).

Figura 6 – A matriz 3 x 3 no centro da figura representa os valores reais de atenuação e os valores das projeções (a, b, c, d) mostram o somatório dos valores de atenuação para este exemplo



Fonte: BUSHBERG, (2021).

No processo de soma das projeções para muitos ângulos, em torno dos 360°, tem como um resultado um maior borramento nas imagens retroprojetadas. É possível melhorar o borramento utilizando filtros matemáticos que são usados nos procedimentos de processamento das imagens (BUSHBERG, 2021).

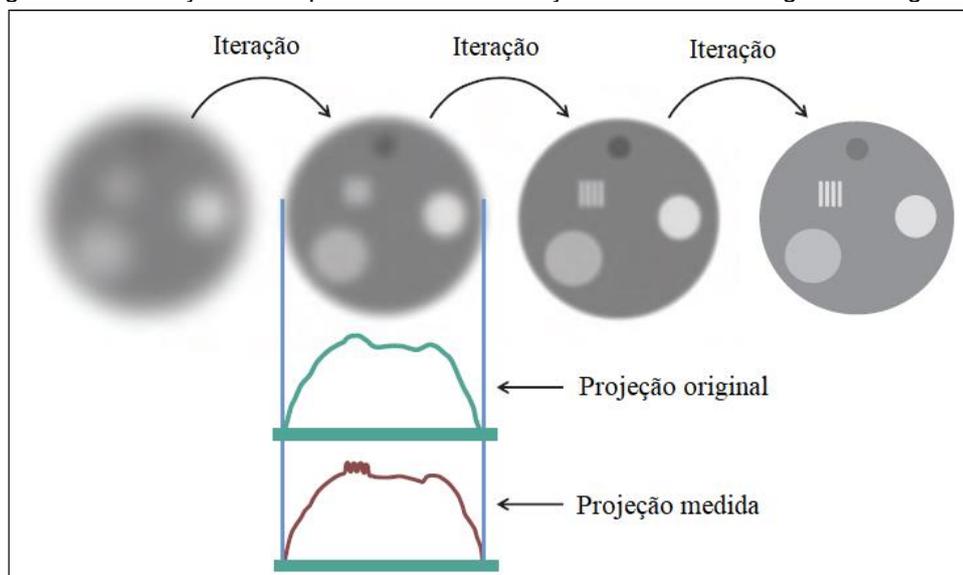
2.1.3.2 Reconstrução iterativa

Os algoritmos da reconstrução iterativa possuem uma maior complexidade computacional em relação a retroprojeção filtrada. Como consequência, a retroprojeção filtrada foi o principal método para reconstruir imagens de TC por muitas décadas. No entanto, com os avanços nos projetos dos algoritmos juntamente aos desenvolvimentos de hardware de computadores mais modernos levaram a algoritmos de reconstrução iterativa clinicamente úteis para imagens tomográficas, mesmo com seu grande conjunto de dados (BUSHBERG, 2021).

A técnica de reconstrução iterativa das imagens tomográficas se dá basicamente por uma série de iterações no processo de reconstrução da imagem, conforme ilustrado na figura 7. A primeira imagem pode começar como uma suposição, ou como

uma imagem de uma retroprojeção filtrada. O processo de iteração passa por uma série de atualizações (iterações), que direcionam a imagem reconstruída a uma representação mais exata do objeto escaneado.

Figura 7 – Ilustração da sequência da reconstrução iterativa de imagens tomográficas



Fonte: Adaptado de BUSHBERG, (2012).

A diferença entre a projeção original e a projeção medida é a matriz de erro, que é calculada para todas as projeções. Cada algoritmo iterativo específico usa a matriz de erro para atualizar a próxima iteração da imagem, com a intenção de reduzir a matriz de erro na iteração subsequente. Existe um equilíbrio entre reconstruir as informações anatômicas e o ruído na imagem. Após um certo número de iterações, a imagem estimada torna-se uma excelente representação do objeto escaneado (BUSHBERG, 2012).

2.2 DIAGNÓSTICO DA COVID-19 POR TOMOGRAFIA COMPUTADORIZADA

Em dezembro de 2019, na cidade de Wuhan, na China, foram descritos alguns casos de infecção pulmonar por um novo coronavírus (SARS-CoV-2). Denominada de COVID-19, a doença logo se disseminou por outras cidades da China e países asiáticos e, posteriormente, por todos os continentes, levando a Organização Mundial da Saúde a decretá-la como pandemia em 11 de março de 2020 (BRASIL, 2020).

A apresentação clínica da COVID-19 pode ser inespecífica, com sintomas comuns a outras síndromes gripais. Os principais sintomas da doença são febre, tosse, fadiga,

anorexia, dispneia, mialgia, dor de garganta, náusea, tonturas, diarreia, cefaleia, vômitos, dor abdominal e anosmia. O período médio de incubação é entre 4 e 5 dias, pode alcançar 14 dias. A maior parte dos pacientes apresenta algum sintoma com até 11,5 dias. O teste considerado padrão-ouro é a reação em cadeia da polimerase por transcriptase reversa (RT-PCR), que consiste na detecção direta do RNA viral em coleta de secreção respiratória da nasofaringe, orofaringe ou dos pulmões (MEIRELLES, 2020).

Os métodos de imagem têm desempenhado um papel proeminente no diagnóstico da COVID-19 e no acompanhamento dos pacientes com a doença. Embora o RT-PCR seja o padrão ouro para o diagnóstico de infecção pela síndrome respiratória aguda grave (SARS-CoV-2), a TC de tórax é uma alternativa para situações em que é necessária a tomada de decisões imediatas. A TC também é usada para investigação do comprometimento da doença e para avaliação de complicações como tromboembolismo pulmonar, infecção bacteriana, abscesso pulmonar, empiema pleural, síndrome da angústia respiratória aguda, miocardite e edema agudo de pulmão (PUGLIA, et al., 2021).

2.2.1 Principais achados radiológicos dos exames de tomografia computadorizada no diagnóstico da COVID-19

O protocolo de exame recomendado para o diagnóstico da COVID-19 por tomografia computadorizada é a TC de alta resolução (TCAR) do tórax. O exame é usualmente realizado sem contraste intravenoso, exceto quando há suspeita de complicações como tromboembolismo pulmonar, sendo então necessário o uso do meio de contraste.

Os principais achados radiológicos nos exames de tomografia no diagnóstico da COVID-19 são opacidades em vidro fosco, pavimentação em mosaico, consolidações, opacidades reticulares, linhas subpleurais, sinal do halo invertido e espessamento pleural. Outros achados são menos comuns, como alterações das vias aéreas, dilatações vasculares, nódulos pulmonares, linfonodomegalias, derrame pleural e derrame pericárdico, estes três últimos usualmente indicando pior prognóstico (MEIRELLES, 2020).

As opacidades em vidro fosco, conforme apresentado na Figura 8, são encontradas na maioria dos pacientes e consistem em manifestação precoce da

doença. São geralmente bilaterais, de predomínio basal periférico, por vezes de aspecto arredondado (MEIRELLES, 2020).

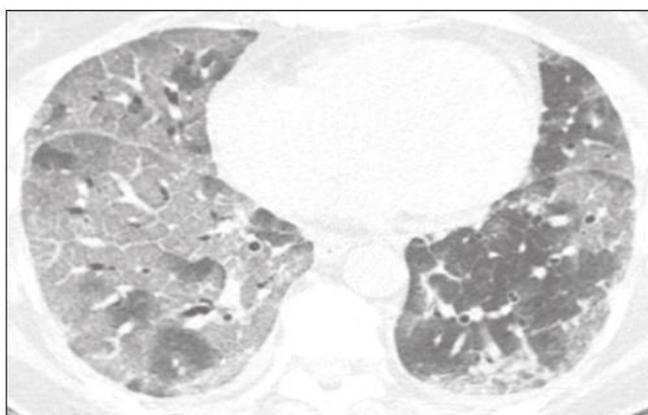
Figura 8 – Tomografia de tórax de paciente do sexo feminino, de 34 anos, com dor torácica e tosse há 4 dias e RT-PCR positivo para SARS-CoV-2. A imagem apresenta opacidades pulmonares com atenuação de vidro fosco, de predomínio periférico, bilaterais



Fonte: MEIRELLES, (2020).

A pavimentação em mosaico, conforme apresentado na Figura 9, é vista na maioria dos pacientes, com maior incidência na fase de pico da doença (cerca de 10 dias), podendo ser decorrente de hiperplasia do interstício intralobular e interlobular pulmonar causada pelo processo inflamatório induzido pela infecção (MEIRELLES, 2020).

Figura 9 – Tomografia de tórax de paciente do sexo masculino, de 52 anos, com anosmia e tosse seca há 6 dias e RT-PCR positivo para SARS-CoV-2. A imagem apresenta padrão pulmonar de pavimentação em mosaico bilateral, com opacidades em vidro fosco associadas a espessamentos septais interlobulares



Fonte: MEIRELLES, (2020).

As consolidações parenquimatosas, conforme apresentado na Figura 10, estão presentes na maioria dos pacientes, principalmente nos pacientes com mais de 60

anos, e indicam fase mais avançada da doença, geralmente surgindo após 10 a 14 dias do início dos sintomas (MEIRELLES, 2020).

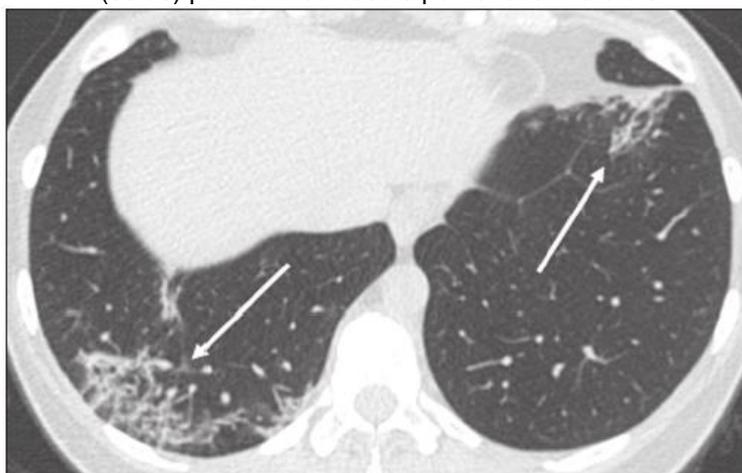
Figura 10 – Tomografia de tórax de paciente do sexo masculino, de 63 anos, com dispneia e episódios de febre há 7 dias, com RT-PCR positivo para SARS-CoV-2. A imagem apresenta consolidações pulmonares (setas) periféricas bilaterais



Fonte: MEIRELLES, (2020).

As opacidades pulmonares reticulares, conforme apresentado na Figura 11, são vistas em alguns pacientes, usualmente em fases mais avançadas da doença e nos indivíduos acima de 60 anos (MEIRELLES, 2020).

Figura 11 – Tomografia de tórax de paciente do sexo feminino, de 71 anos, com dispneia e tosse seca há 8 dias, com RT-PCR positivo para SARS-CoV-2. A imagem apresenta opacidades reticulares (setas) pulmonares basais periféricas bilaterais



Fonte: MEIRELLES, (2020).

As linhas subpleurais, conforme apresentado na Figura 12, são menos comuns nos pacientes, podendo representar edema pulmonar ou evolução para fibrose. São mais comuns em pacientes acima dos 60 anos e após 10 dias de doença (MEIRELLES, 2020).

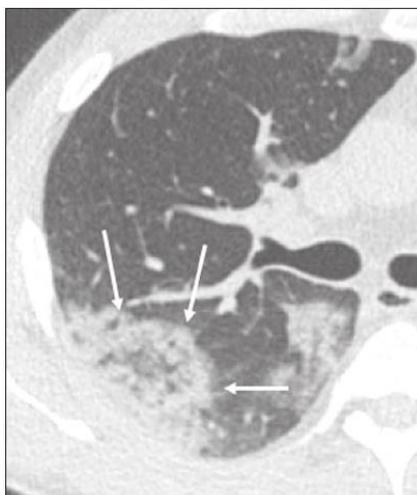
Figura 12 – Tomografia de tórax mostrando linha subpleural (setas) em paciente com diagnóstico confirmado de COVID-19, com tosse seca, dispneia e anosmia há 9 dias



Fonte: MEIRELLES, (2020).

O sinal do halo invertido, conforme Figura 13, é geralmente visto em fases mais tardias da doença e é menos frequente nos pacientes, pode ser decorrente de pneumonia em organização ou infartos pulmonares associados (MEIRELLES, 2020).

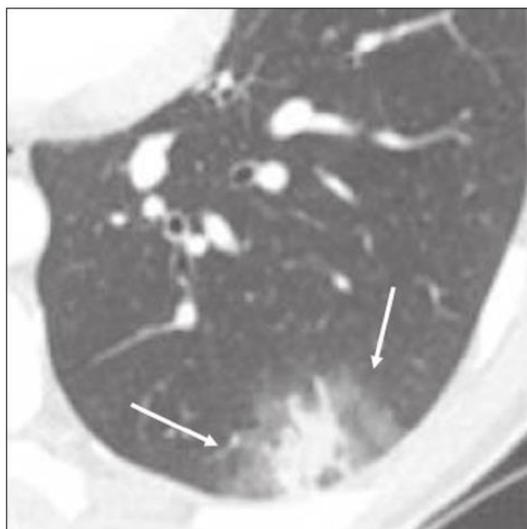
Figura 13 – Tomografia de tórax de paciente do sexo masculino, de 59 anos, internado na unidade de terapia intensiva há 12 dias com diagnóstico de COVID-19. A imagem apresenta opacidades pulmonares, uma delas com sinal do halo invertido (setas), caracterizado por consolidação com centro de baixa atenuação



Fonte: MEIRELLES, (2020).

O sinal do halo, com atenuação de vidro fosco ao redor de focos de consolidação, conforme apresentado na Figura 14, é infrequente. O diagnóstico pode ser confundido com outras infecções, vasculites e neoplasias. (MEIRELLES, 2020).

Figura 14 – Tomografia de tórax mostrando sinal do halo caracterizado por consolidação nodular pulmonar com halo de vidro fosco (setas) em paciente com diagnóstico de COVID-19



Fonte: MEIRELLES, (2020).

Alterações das vias aéreas, como broncogramas aéreos, são infrequentes e podem indicar gravidade. Espessamento brônquico, bronquiectasias e nódulos centrolobulares são raros em adultos, sendo mais encontrados na população pediátrica (MEIRELLES, 2020).

2.2.2 Desafios no diagnóstico da COVID-19 por tomografia computadorizada

Em tempos de pandemia mundial de uma doença com relativa facilidade propagação entre as pessoas, o diagnóstico diferencial é fundamental para o controle da doença e para o correto tratamento dos pacientes. Os sintomas clínicos e radiológicos da COVID-19 podem ser confundidos com os sintomas de outras doenças, incluindo gripes, pneumonias, síndrome respiratória aguda grave (SARS), síndrome respiratória do Oriente Médio (MERS). Atualmente a avaliação clínica do paciente e de exames de imagens podem indicar a suspeita da COVID-19, sendo o teste de reação em cadeia da polimerase por transcriptase reversa (RT-PCR) como o mais indicado para confirmação do diagnóstico (ZU et al., 2020).

Entretanto, o diagnóstico da COVID-19 por meio de exames de TC de tórax pode ser otimizado. Soluções de Inteligência Artificial (IA) e Radiômica baseadas em exames de TC de tórax são fundamentais na detecção, diagnóstico e quantificação dos achados pulmonares decorrentes da infecção viral e para predição de complicações da doença, com o intuito de se instituir um correto tratamento (FANG et al., 2020). O maior benefício potencial de sistemas computacionais no auxílio diagnóstico é a triagem rápida dos pacientes com suspeita clínica da doença, principalmente em locais mais afetados pela pandemia, com muitos casos e poucos recursos diagnósticos.

2.3 RADIÔMICA

O sufixo “ômica” é um termo originado na biologia molecular para descrever a caracterização detalhada de moléculas como o DNA (genômica) e as proteínas (proteômica). A radiômica é definida pela extração massiva de dados mensuráveis de imagens médicas (*features* ou atributos radiômicos) e sua integração em modelos preditivos multidisciplinares para o manejo diagnóstico, terapêutico e prognóstico de pacientes (KUMAR et al., 2012).

A radiômica pode ser usada para capturar propriedades de tecidos e lesões. Por exemplo, em oncologia, a avaliação da heterogeneidade do tecido é relevante, pois análises genômicas demonstraram que o grau de heterogeneidade do tumor é um determinante prognóstico de sobrevivência e um obstáculo no controle de câncer. Enquanto as biópsias capturam a heterogeneidade em apenas uma pequena porção de um tumor e geralmente em apenas um único sítio anatômico, a radiômica captura a heterogeneidade em todo o volume do tumor. Além disso, os atributos radiômicos também estão associados à agressividade do tumor. Os atributos radiômicos foram sugeridos para prever desfechos clínicos, como sobrevivência e resposta ao tratamento.

Embora até mesmo os atributos radiômicos individuais possam se correlacionar com dados genômicos ou resultados clínicos, o impacto da radiômica aumenta quando a riqueza de informações que ela fornece (normalmente centenas de atributos), é processada usando técnicas de aprendizado de máquina. Além disso, os dados radiômicos podem ser minerados, o que significa que um conjunto de dados suficientemente grande pode ser usado para descobrir marcadores e padrões

previamente desconhecidos de evolução, progressão e resposta ao tratamento da doença (MAYERHOEFER et al., 2020).

Em face dos recentes avanços nas terapias-alvo e imunoterapias, principalmente no tratamento das neoplasias, a necessidade de uma abordagem de análise robusta de imagens médicas se tornou imperativa e a radiômica tem potencial para prover isso de maneira não invasiva, rápida e de baixo custo (SANTOS et al., 2019).

2.3.1 Atributos radiômicos

Uma vez que as regiões de interesse são definidas, os atributos radiômicos podem ser extraídos das imagens médicas. Os atributos radiômicos podem ser subdivididos estatisticamente, incluindo atributos baseados em histograma, em textura, em modelo, em transformadas e baseados na forma (MAYERHOEFER et al., 2020).

De acordo Machado e colaboradores (2021), a extração de atributos radiômicos fornece um método poderoso para avaliar as diferenças na biologia do tumor, identificando preditores que podem ter um papel funcional em fenótipos específicos.

2.3.1.1 Atributos baseados em histogramas

Os atributos estatísticos mais simples são baseados no histograma global de nível de cinza e incluem média, máximo, mínimo, variação e percentis do nível de cinza. Como esses recursos são baseados em análises de pixel único ou voxel único, eles são chamados de recursos de primeira ordem. Atributos mais sofisticados incluem assimetria e curtose, que descrevem a forma da distribuição de intensidade dos dados: a assimetria reflete a assimetria da curva de distribuição de dados para a esquerda (inclinação negativa, abaixo da média) ou direita (inclinação positiva, acima da média), enquanto a curtose reflete a cauda de uma distribuição de dados em relação a uma distribuição gaussiana. Outras características incluem entropia e uniformidade do histograma (MAYERHOEFER et al., 2020).

2.3.1.2 Atributos baseados em textura

Os atributos radiômicos baseados em textura podem ser subdivididos em classes. Uma abordagem simples para a descrição da textura radiômica é a análise do

gradiente absoluto, que reflete o grau das flutuações de intensidade do nível de cinza em uma imagem. Por exemplo, para dois pixels adjacentes ou voxels, o gradiente é mais alto se um for preto e o outro branco, enquanto, se ambos os pixels forem pretos (ou ambos forem brancos) o gradiente nessa localização será zero. Se o nível de cinza aumenta de preto para branco, o gradiente é positivo; se diminui de branco para preto, o gradiente é negativo. Semelhante aos atributos de histograma, os atributos de gradiente incluem gradiente médio, variância, assimetria e curtose.

Outra classe de atributo baseado em textura é a matriz de coocorrência de nível de cinza – GLCM (do inglês - *Gray-Level Cooccurrence Matrix*). O GLCM é um histograma de nível de cinza de segunda ordem que captura relações espaciais de pares de pixels ou voxels com intensidades de nível de cinza predefinidas, em diferentes direções (horizontal, vertical ou diagonal para uma análise 2D ou 13 direções para uma análise 3D) e com uma distância predefinida entre os pixels ou voxels. A classe de atributos do GLCM inclui entropia (uma medida de não homogeneidade ou aleatoriedade do nível de cinza), segundo momento angular (também chamado de uniformidade ou energia que reflete a homogeneidade ou ordem do nível de cinza) e contraste (que enfatiza as diferenças de nível de cinza entre pixels ou voxels pertencentes a um par de pixels ou voxels).

A matriz de comprimento de execução de nível de cinza – GLRLM (do inglês - *Gray-Level Run-length Matrix*) é a classe de atributo baseado em textura que fornece informações sobre a distribuição espacial de execuções de pixels consecutivos com o mesmo nível de cinza, em uma ou mais direções, em 2 ou 3 dimensões. A classe de atributos do GLRLM inclui fração, que avalia a porcentagem de pixels ou voxels dentro do volume segmentado que faz parte das execuções e, portanto, reflete a granulação.

Baseado no mesmo princípio do GLRLM, a classe de atributos matriz de tamanho de zona de nível de cinza – GLSZM (do inglês - *Gray-Level Size Zone Matrix*) em que as contagens do número de grupos de pixels ou voxels vizinhos interconectados com o mesmo nível de cinza (as chamadas zonas) formam a base para a matriz. Uma textura mais homogênea resultará em uma matriz mais ampla e plana. O GLSZM não é calculado para diferentes direções, mas pode ser calculado para diferentes pixels ou distâncias de voxel que definem a vizinhança. Os atributos da classe GLSZM podem ser calculados em 2 dimensões (8 pixels vizinhos) ou 3 dimensões (26 voxels vizinhos) e, seguindo as definições do GLRLM, incluem fração (porcentagem de pixels ou voxels que fazem parte das zonas).

Como uma variação do GLSZM, a classe de atributo matriz de distância de zona de nível cinza – GLDZM (do inglês - *Gray-Level Distance Zone Matrix*) não apenas avalia as zonas de pixels ou voxels vizinhos interconectados com o mesmo nível de cinza, mas exige que eles estejam à mesma distância da borda do volume segmentado.

A classe de atributo baseado em textura de matriz de diferença de tons de cinza da vizinhança – NGTDM (do inglês - *Neighborhood Gray-Tone Difference Matrix*) quantifica a soma das diferenças entre o nível de cinza de um pixel ou voxel e o nível de cinza médio de seus pixels ou voxels vizinhos dentro de uma distância predefinida. Os principais atributos NGTDM incluem aspereza (do inglês – *coarseness*), ocupação (do inglês – *busyness*) e complexidade (do inglês – *complexity*). A aspereza reflete a diferença de nível de cinza entre o pixel central ou voxel e sua vizinhança e, portanto, captura a taxa espacial de mudanças nas intensidades de nível de cinza, ou seja, um volume segmentado que consiste em volumes maiores com níveis de cinza relativamente uniformes (uma taxa mais baixa de mudanças de intensidade espacial) terá um alto valor de aspereza. A ocupação, por outro lado, reflete mudanças rápidas de nível de cinza entre o pixel central ou voxel e seus vizinhos (ou seja, uma alta frequência espacial de mudanças de intensidade), de modo que um volume segmentado compreendendo muitas áreas pequenas com níveis de cinza marcadamente diferentes terá uma maior ocupação.

A classe de atributo radiômico baseado em textura matriz de dependência de nível de cinza da vizinhança – NGLDM (do inglês - *Neighborhood Gray-Level Dependence Matrix*) também é baseado na relação de nível de cinza entre um pixel ou voxel central e sua vizinhança. Sendo que no NGLDM, um pixel ou voxel vizinho dentro de uma distância predefinida é considerado como conectado ao pixel ou voxel central se atender ao critério de dependência em termos de uma faixa definida de diferenças de nível de cinza (MAYERHOEFER et al., 2020).

2.3.1.3 Atributos baseados em modelos

As análises baseadas em modelos visam interpretar informações espaciais de nível de cinza para caracterizar objetos ou formas. Um modelo parametrizado de geração de textura é calculado e ajustado ao volume segmentado, e seus parâmetros estimados são usados como atributos radiômicas. O modelo autoregressivo é um

exemplo de abordagem baseada em modelo e é fundamentado na ideia de que o nível de cinza de um pixel é uma soma ponderada dos níveis de cinza de 4 pixels vizinhos.

A análise fractal também produz atributos que podem ser usados para radiômica, em particular dimensão fractal, que reflete a taxa de adição de detalhes estruturais com aumento, escala ou resolução e, portanto, serve como uma medida de complexidade (MAYERHOEFER et al., 2020).

2.3.1.4 Atributos baseados em formas

Os atributos baseados em forma descrevem propriedades geométricas dos volumes segmentados. Muitos atributos baseados em forma são conceitualmente muito mais simples do que outros recursos radiômicos, como diâmetros 2D e 3D, eixos e suas proporções. Abordagens baseadas em superfície e volume baseadas no uso de malhas (isto é, pequenos polígonos como triângulos e tetraedros) são mais complexas. Os recursos incluem compactação e esfericidade, que descrevem, por exemplo, como a forma de um volume segmentado difere de um círculo (para análises 2D) ou de uma esfera (para análises 3D) (MAYERHOEFER et al., 2020).

2.3.1.5 Atributos baseados em transformadas

Os métodos baseados em transformadas, incluindo Fourier, Gabor e Haar Wavelet, analisam os padrões de nível de cinza em um espaço diferente. A transformada Haar Wavelet, por exemplo, analisa o conteúdo de frequência de uma imagem em diferentes escalas. A decomposição Wavelet de uma imagem é possível aplicando um par de filtros de espelho de quadratura, um filtro passa-alta e um filtro passa-baixa. Embora o filtro passa-alta destaque as alterações no nível de cinza e, portanto, enfatize os detalhes da imagem, o filtro passa-baixa suaviza a imagem em termos de nível de cinza, removendo os detalhes da imagem. Após a decomposição do sinal, um conjunto de canais de frequência é espacialmente orientado, o qual é usado para descrever a variabilidade local da imagem. As energias dentro dos canais de frequência são então usadas como atributos radiômicos. (MAYERHOEFER et al., 2020).

2.3.1.6 Seleção (qualificação) dos atributos radiômicos

Uma vez que a extração dos atributos radiômicos for concluída, os atributos relevantes que serão usados no modelo estatístico para resolver o problema clínico (por exemplo, para distinguir entre lesões benignas e malignas) devem ser identificados (MAYERHOEFER et al., 2020).

Centenas de atributos radiômicos podem ser calculados. No entanto, todos esses atributos extraídos podem não ser úteis para uma tarefa específica. Além disso, o número de atributos extraídos pode ser maior do que o número de amostras em um estudo, reduzindo o poder e aumentando a probabilidade de *overfitting* dos dados. Portanto, a redução da dimensionalidade e a seleção de atributos específicos da tarefa para melhor desempenho são etapas necessárias. Diferentes métodos de seleção de atributos podem ser usados para esse propósito e podem explorar o aprendizado de máquina ou abordagens estatísticas (KUMAR et al., 2012).

2.3.2 Harmonização dos atributos radiômicos

A harmonização é uma técnica de pós-processamento matemático para remover o chamado efeito de lote (do inglês - *batch effect*) em características radiômicas após a aquisição, a reconstrução e a análise de imagem (efeitos dependentes de variações de parâmetros de diferentes centros de aquisição). A harmonização, portanto, não é aplicada a imagens, mas diretamente aos valores dos atributos radiômicos (MAYERHOEFER et al., 2020).

Os atributos radiômicos são sensíveis aos parâmetros de aquisição e reconstrução das imagens tomográficas, devido aos diferentes filtros, protocolos, modelos de equipamentos etc. Os atributos, no entanto, não podem ser diretamente comparáveis entre imagens adquiridas em diferentes equipamentos, ou diferentes protocolos, o que pode limitar seu uso em estudos multicêntricos. Uma alternativa para reduzir as alterações dos atributos radiômicos obtidos em estudos multicêntricos é a aplicação de métodos de harmonização. Segundo Machado e colaboradores (2019), o uso de métricas harmonizadas pode ajudar na medicina personalizada, uma vez que permitem uma avaliação mais confiável das características do tumor.

Existem vários métodos de harmonização de dados, como *global scaling*, *functional normalization*, RAVEL, *Surrogate Variable Analysis* (SVA) e ComBat. Entretanto, de acordo com Orhac e colaboradores (2019), o método de harmonização ComBat é o mais eficiente na correção dos atributos radiômicos adquiridos em

estudos multicêntricos. O método de harmonização ComBat, desenvolvido por Johnson e colaboradores (2007), é baseado em dados que identifica o efeito da variação entre as aquisições dos exames em diferentes centros de imagens i assumindo o valor de cada atributo y_{ijv} medido em um volume de interesse v com o equipamento j :

$$y_{ijv} = \alpha_v + X_{ij}\beta_v + \gamma_{iv} + \delta_{iv}\varepsilon_{ijv} \quad \text{Eq. 3}$$

onde α_v é o valor médio global do atributo para um volume de interesse v , X é uma matriz de design para as covariantes de interesse (por exemplo, gênero, idade etc.), β_v é um vetor específico da região de interesse v de coeficientes de regressão correspondentes a X , os parâmetros γ_{iv} e δ_{iv} correspondem, respectivamente, aos efeitos aditivos e multiplicativos para o centro de imagem i e para a região de interesse v e ε_{ijv} é um termo de erro que segue uma distribuição normal com média zero e variância σ_v^2 .

O método ComBat usa uma estrutura empírica de Bayes para melhorar a variância dos parâmetros, estimando uma distribuição estatística empírica para cada parâmetro, assumindo que todas as regiões de interesse compartilham uma distribuição comum. Nesse sentido, as informações de todas as regiões de interesse são utilizadas para informar as propriedades estatísticas dos efeitos ocasionados pelos diferentes centros de imagens.

A compensação consiste em estimar os parâmetros do modelo α_v , β_v , γ_{iv} e δ_{iv} usando uma abordagem de máxima verossimilhança com base no conjunto de observações disponíveis:

$$y_{ijv}^{ComBat} = \frac{y_{ijv} - \hat{\alpha}_v - X_{ij}\hat{\beta}_v - \hat{\gamma}_{iv}}{\hat{\delta}_{iv}} + \hat{\alpha}_v + X_{ij}\hat{\beta}_v \quad \text{Eq. 4}$$

Em que $\hat{\alpha}_v$, $\hat{\beta}_v$, $\hat{\gamma}_{iv}$ e $\hat{\delta}_{iv}$ são dados estimados dos parâmetros α_v , β_v , γ_{iv} e δ_{iv} , respectivamente (JOHNSON et al., 2007). A harmonização ComBat determina uma transformação para cada atributo separadamente, baseado no efeito da variação de cada centro de imagem observado nos valores dos atributos (ORLHAC et al., 2019).

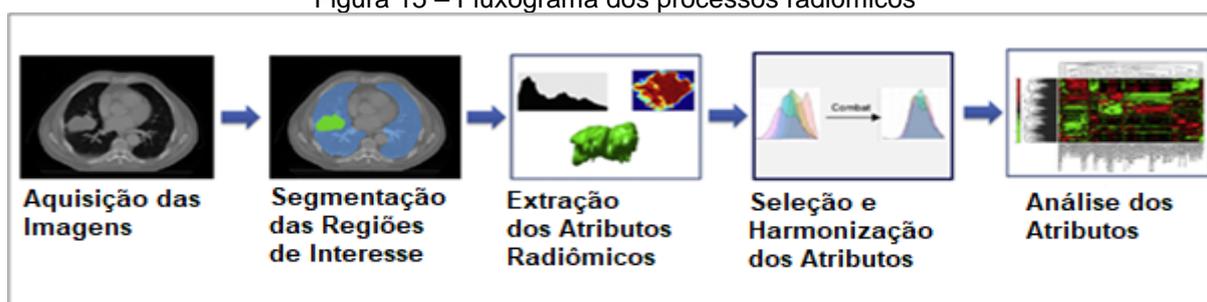
O método de harmonização ComBat oferece uma solução para realinhar os atributos radiômicos com várias vantagens. O ComBat é facilmente disponível, não há a necessidade de realizar novas aquisições de imagens, não há a necessidade de um conjunto de aprendizagem, não altera as definições dos atributos e pode ser utilizado com qualquer atributo radiômico.

3 METODOLOGIA

Este trabalho integra um projeto de pesquisa sob a coordenação do Laboratório de Pesquisa em Infectologia (LAPI) do HUPES/UFBA, com aprovação no Comitê de Ética em Pesquisa (CAAE: 36624620.0.0000.0049), com o título “Desenvolvimento e validação de software para diagnóstico de COVID-19 baseado em Tomografia Computadorizada, Radiômica e Inteligência Artificial”. O projeto de pesquisa visa disponibilizar um produto para diagnóstico do COVID-19, utilizando imagens de TC de tórax e inteligência artificial (IA).

O primeiro passo do fluxograma de aplicação radiômica foi a aquisição das imagens médicas analisadas. A partir das imagens, as regiões de interesse foram definidas por meio das segmentações manuais, que foram utilizadas para treinar uma rede neural de segmentação automática. Em seguida, os atributos radiômicos foram extraídos dos volumes segmentados automaticamente. Por ser um estudo radiômico multicêntrico, com diferentes equipamentos de tomografia computadorizada e diferentes protocolos de aquisição e reconstrução das imagens, os atributos foram selecionados, analisados e harmonizados. As etapas do trabalho são apresentadas no fluxograma da Figura 15.

Figura 15 – Fluxograma dos processos radiômicos



Fonte: Adaptado de LAMBIN *et al.*, (2012).

3.1 AQUISIÇÃO DAS IMAGENS

Os exames de TC de tórax dos pacientes foram adquiridos em três centros de imagens com os equipamentos apresentados na Tabela 1.

- Complexo Hospitalar Universitário Prof. Edgard Santos – HUPES/UFBA (Hospital A);
- Hospital Universitário Alcides Carneiro – HUAC/UFCG (Hospital B); e

- Hospital Universitário de Juiz de Fora – HUJF/UFJF (Hospital C).

A pesquisa foi aprovada pelo comitê de ética de cada instituição, conforme apresentado no ANEXO A - PARECER CONSUBSTANCIADO DO CEP: Hospital Universitário Prof. Edgard Santos (UFBA) (4.231.669), Hospital Universitário Alcides Carneiro (UFMG) (4.818.271) e Hospital Universitário de Juiz de Fora (UFJF) (4.926.688).

Tabela 1 – Características dos equipamentos

Hospital	Hospital A	Hospital B	Hospital C
Fabricante do Tomógrafo	Toshiba	Philips	Siemens
Modelo do Tomógrafo	Aquilion	Brilliance 6	Emotion 6

Fonte: O autor (2022).

Os pesquisadores de cada centro de imagem selecionaram retrospectivamente os exames de TC de tórax relacionada a COVID-19 realizados entre abril de 2020 e abril de 2021, com confirmação da COVID-19 no teste RT-PCR. Além disso, os pesquisadores de cada centro também selecionaram retrospectivamente os exames de TC de tórax relacionados a pneumonias (casos controle) antes da pandemia de COVID-19, realizados entre janeiro de 2018 e outubro de 2019, conforme apresentado na Tabela 2.

Os critérios de exclusão foram: imagens tomográficas com má qualidade; lesões pequenas ou imperceptíveis; indisponibilidade do teste RT-PCT para o grupo COVID-19; imagens que não estavam em conformidade com o protocolo institucional para imagens de pulmão (por exemplo, tamanho da matriz, kernel) e imagens que não estavam mais disponíveis no PACS (do inglês - *Picture Archiving and Communication System*).

Tabela 2 – Características dos exames obtidos

Hospital	Hospital A	Hospital B	Hospital C
Pacientes COVID-19 / Controle	14/ 10	14/ 16	09/ 10
Média de Idade	57,0 anos	57,8 anos	58,1 anos
Comprometimento Pulmonar			

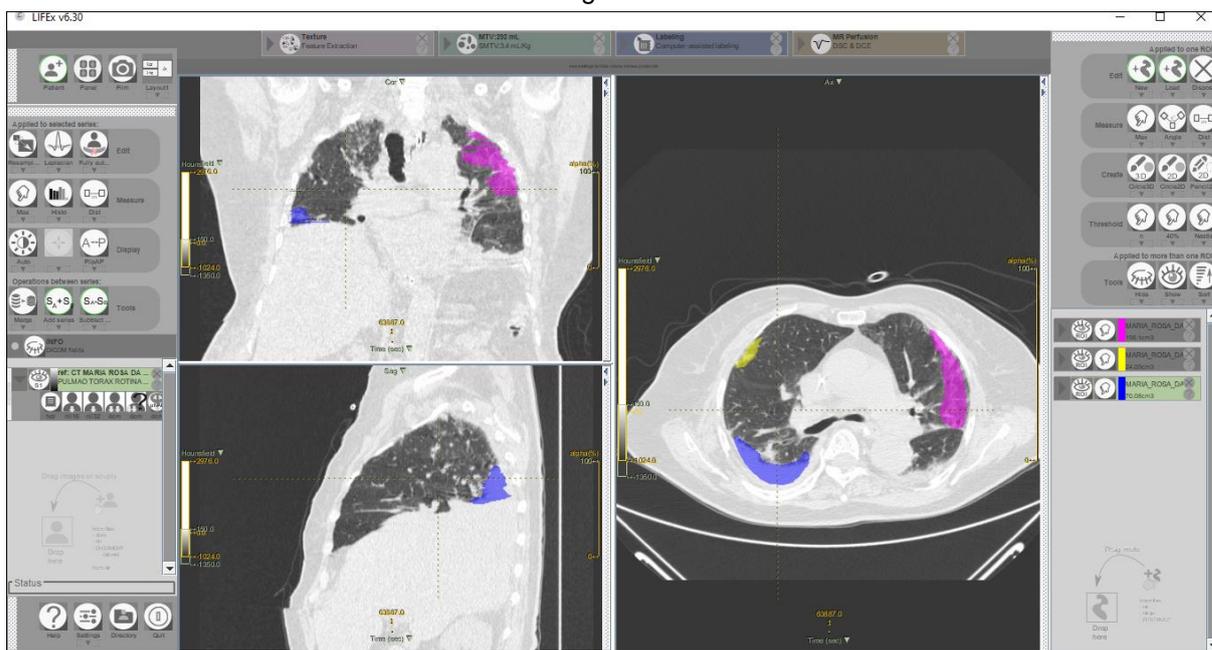
Baixo (< 25%)	14	10	07
Moderado (25 - 50%)	04	10	08
Severo (> 50%)	06	10	04

Fonte: O autor (2022).

3.2 SEGMENTAÇÃO DAS REGIÕES DE INTERESSE E EXTRAÇÃO DOS ATRIBUTOS RADIÔMICOS

Os médicos radiologistas de cada centro de imagem segmentaram manualmente as lesões pulmonares relacionadas a pneumonias dos seus pacientes. As regiões de interesse - ROI (do inglês - *region of interest*) nas imagens foram segmentadas utilizando o software Lifex, versão 6.30 (www.lifexsoft.org). A Figura 16 apresenta uma ilustração da interface de segmentação de regiões de interesse do Lifex.

Figura 16 – Layout da interface do Lifex ilustrando três segmentações de regiões de interesse em uma tomografia de tórax



Fonte: O autor (2022).

As segmentações manuais realizadas pelos médicos radiologistas de cada centro de imagem utilizando o Lifex foram usadas previamente para treinar uma rede neural convolucional 2D com um esquema de validação cruzada de 5 vezes. Esta rede neural foi desenvolvida pelos pesquisadores do grupo de pesquisa para segmentação automática das lesões, onde cada conjunto com 60 pacientes foi usado para o

treinamento da rede, e os pacientes restantes foram usados para segmentação e extração dos atributos radiômicos. Sendo que, foram utilizados 12.870 cortes tomográficos, e os dados foram aumentados, realizando erosão de 1 mm, rotação e translação nas ROI's segmentadas manualmente, resultando em um total de 51.120 cortes. Os mapas de probabilidade de classe 2D de saída foram filtrados com um kernel gaussiano 3D antes do limiar para obter a segmentação final da lesão 3D. Apenas as lesões segmentadas independentemente foram incluídas na análise de seleção dos atributos radiômicos. Ao todo foram 1060 lesões segmentadas, sendo 446 lesões no Hospital A, 234 no Hospital B e 380 no Hospital C.

Os atributos foram extraídos destas segmentações utilizando o PyRadiomics, uma plataforma de código aberto que permite o processamento e a extração de atributos radiômicos de imagens médicas (GRIETHUYSEN, VAN et al., 2017). A grade da matriz foi reamostrada para voxels de $1 \times 1 \times 2 \text{ mm}^3$ e a discretização dentro de cada volume de interesse – VOI (do inglês - *volume of interest*) foi dimensionada para 128 níveis de tons de cinza. Os atributos extraídos das segmentações automáticas foram utilizados neste trabalho.

3.3 SELEÇÃO E HARMONIZAÇÃO DOS ATRIBUTOS RADIÔMICOS

Para a seleção dos atributos avaliados neste trabalho, foram realizados dois métodos de seleção: um método não supervisionado e um método supervisionado.

3.3.1 Método não supervisionado de seleção

No método não supervisionado para seleção dos atributos radiômicos, foram analisadas as reprodutibilidades dos atributos com relação a variações devido a diferentes segmentadores, protocolos e equipamentos nas aquisições dos exames, conforme descrito abaixo:

- Para simular diferentes segmentadores, foram realizadas erosões de 1 mm e 3 mm nos VOI's segmentados automaticamente, resultando em três conjuntos de imagens segmentadas (imagem original, com 1 mm de erosão e com 3 mm de erosão); e
- Para simular os diferentes protocolos e equipamentos de aquisição, foram aplicados filtros kernels de borramento e de melhoramento de bordas nas

imagens, resultando em três conjuntos de imagens segmentadas (imagem original, com filtro kernels de borramento e com filtro kernels de melhoramento de bordas).

A reprodutibilidade do atributo foi estimada usando o coeficiente de correlação intraclasse – ICC (do inglês - *intraclass correlation coefficient*), que avaliou a concordância de quantificação radiômica para a simulação de efeitos devido aos diferentes segmentadores (imagens com erosões nos VOI's) e para a simulação de efeitos devido aos diferentes protocolos e equipamentos de aquisição (imagens com filtros). Os atributos com ICC > 0,90 concomitantemente nas avaliações com erosões nas VOI's e com filtros nas imagens foram selecionados e considerados mais robustos. O cálculo do ICC foi realizado utilizando o código Python disponível em <https://github.com/Mind-the-Pineapple/ICC.git>, utilizando o Python-3.9.2.

3.3.2 Método supervisionado de seleção

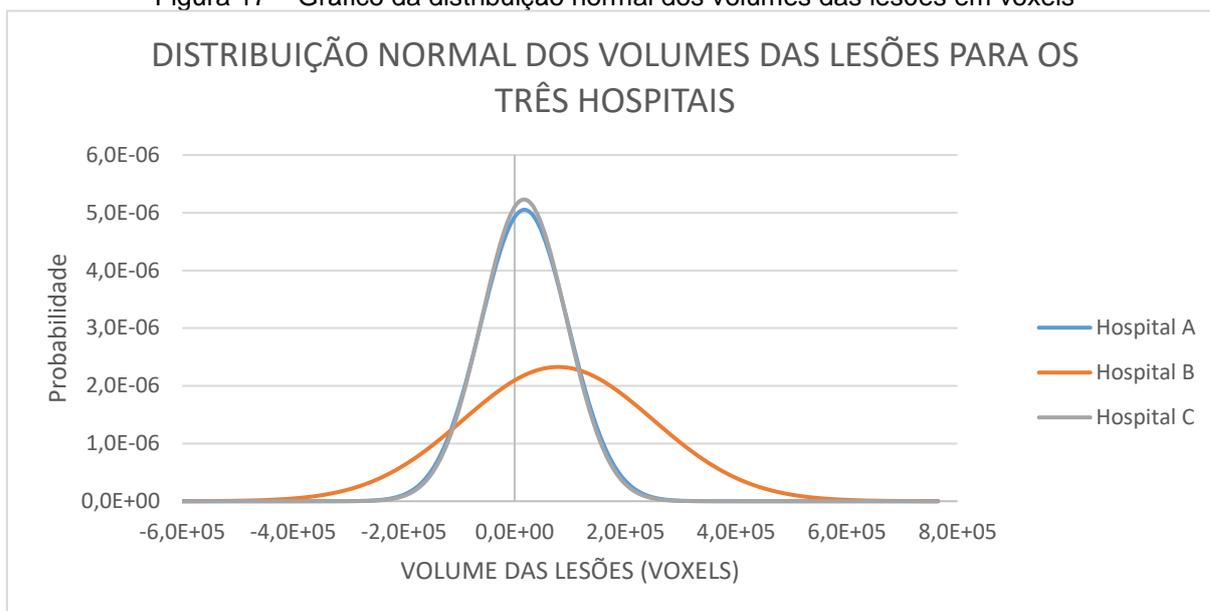
Já no método supervisionado para seleção dos atributos radiômicos, o poder discriminatório do biomarcador em relação ao status da lesão (COVID-19 ou não-COVID-19) foi calculado por meio da área sobre a curva ROC (AUC). Os atributos com os menores *p-value* são considerados melhores na predição da doença. Desta forma, para este método, foram selecionados os atributos de cada classe com o menor *p-value*. O cálculo do *p-value* foi realizado utilizando o código Python disponível em <https://github.com/Mind-the-Pineapple/ICC.git>, utilizando o Python-3.9.2.

Partindo do pressuposto que um paciente infectado com COVID-19 apresente mais VOI's com características típicas de COVID-19 e que pacientes com pneumonias não associadas ao COVID-19 apresente menos VOI's associados ao COVID-19, então foi atribuído a cada VOI o mesmo status do paciente (COVID-19 ou não-COVID-19).

3.3.3 Avaliação da correlação do atributo com o volume da lesão

As lesões que foram segmentadas automaticamente nos exames dos pacientes dos três hospitais foram analisadas. Um gráfico da distribuição dos volumes das lesões (com ajuste para distribuição normal), foi plotado para cada hospital, conforme apresentado na Figura 17.

Figura 17 – Gráfico da distribuição normal dos volumes das lesões em voxels



Fonte: O autor (2022).

Por haver uma elevada dispersão entre os volumes das lesões dos pacientes do Hospital B em relação aos outros dois hospitais, foi realizada a avaliação da correlação do atributo radiômico selecionado com o volume da lesão por meio da correlação de Spearman (ρ). Essa correlação indica o quanto o atributo está relacionado ao volume da lesão (para valores de ρ próximos de -1 ou 1), ou o quanto o atributo tem baixa relação com o volume da lesão (para valores de ρ próximos de 0). Segundo Mukaka (2012), para valores de $\rho > 0,7$ ou $\rho < -0,7$ a correlação de Spearman pode ser considerada fortemente relacionada. O cálculo do coeficiente da correlação de Spearman foi realizado utilizando o código Python disponível em <https://github.com/matthijsz/weightedcorr.git>, utilizando o Python-3.9.2.

Para este trabalho, os atributos radiômicos que obtiveram $\rho > 0,7$ ou $\rho < -0,7$ foram considerados fortemente correlacionados com o volume da lesão. Além disso, foram plotados os gráficos de dispersões entre o atributo normalizado e o volume normalizado da lesão. As normalizações foram realizadas pelo cálculo do ranking médio (ranking de Besson) entre o atributo e o volume da lesão.

3.3.4 Método de harmonização ComBat

Os atributos radiômicos foram extraídos sem a aplicação de filtros ou tratamento nas imagens de TC. Um total de 94 atributos foram extraídos para cada lesão,

incluindo 18 atributos de primeira ordem e 76 atributos de textura, conforme apresentado na lista de atributos radiômicos do Apêndice B. Os atributos selecionados e considerados fracamente correlacionados com o volume da lesão foram harmonizados utilizando o método ComBat de harmonização. Esse método foi desenvolvido por Johnson e colaboradores (2007) e tem sua fundamentação matemática descrita nas Eq. 3 e Eq. 4. O método de harmonização ComBat foi aplicado utilizando a biblioteca NeuroCombat disponível em Python (<https://github.com/jmyoung36/neuroCombat.git>).

A harmonização ComBat dos atributos radiômicos foi aplicada com o código Python desenvolvido pelo grupo de pesquisa para este trabalho, conforme apresentado no Apêndice A, utilizando o Python-3.9.2. Para este código, a biblioteca NeuroCombat foi importada para ser utilizada no cálculo, além disso, foram declarados o número de atributos radiômicos e o arquivo CSV com os atributos radiômicos brutos (sem harmonização). Os três hospitais foram declarados no código, sendo que um hospital deve ser declarado como centro de referência para a harmonização. Foram realizadas várias harmonizações variando os centros de referência; essa variação nos centros de referência é necessária para analisar, para cada centro, a melhor harmonização que pode ser utilizada no estudo. Por este motivo, o centro de referência que foi selecionado para este trabalho foi o do Hospital A. Um arquivo CSV com os atributos radiômicos harmonizados é a saída do código apresentado no Apêndice A.

4 RESULTADOS E DISCUSSÃO

4.1 SELEÇÃO DOS ATRIBUTOS RADIÔMICOS

A seleção dos atributos foi realizada por dois métodos, sendo um não supervisionado e outro supervisionado. Dos atributos extraídos, três foram considerados robustos pelo método não supervisionado (com ICC > 0,90 nas avaliações com erosão e com imagens com filtros), conforme apresentados no Apêndice B.

No método supervisionado, 38 atributos obtiveram um *p-value* < 0,05, conforme apresentado no Apêndice B. No entanto, foi selecionado um atributo de cada classe com o menor *p-value*. Desta forma, foram selecionados seis atributos radiômicos pelo método supervisionado para este trabalho.

Além disso, para os nove atributos selecionados foram realizadas as avaliações de correlação de Spearman (ρ) onde é avaliada a correlação entre o atributo radiômico e o volume da lesão.

Os nove atributos radiômicos selecionados, os métodos de seleção, os valores de ICC nas avaliações da erosão dos VOI's e dos filtros das imagens, os *p-value* dos atributos e a suas correlações ρ de Spearman com o volume da lesão são apresentados na Tabela 3.

Tabela 3 – Atributos selecionados

Atributo	Nome	Método de Seleção	ICC – Erosão / Filtro	p-value	ρ
01	GLDM GRAY LEVEL NON UNIFORMITY	Não supervisionado	0,903 / 0,951	0,098	0,917
02	FIRSTORDER ENERGY	Não supervisionado	0,907 / 0,999	0,739	0,837
03	FIRSTORDER TOTAL ENERGY	Não supervisionado	0,907 / 0,999	0,739	0,837
04	FIRSTORDER MEAN	Supervisionado	0,502 / 0,998	< 0,001	0,030
05	GLCM INVERSE VARIANCE	Supervisionado	0,582 / 0,150	< 0,001	-0,030
06	GLDM DEPENDENCE NON	Supervisionado	0,151 / 0,480	< 0,001	-0,780

07	UNIFORMITY NORMALIZED GLRLM LONG RUN LOW GRAY LEVEL EMPHASIS	Supervisionado	0,516 / 0,075	0,001	-0,312
08	GLSZM SMALL AREA EMPHASIS	Supervisionado	0,417 / 0,262	0,013	0,615
09	NGTDM BUSYNESS	Supervisionado	0,620 / 0,072	< 0,001	0,576

Fonte: O autor (2022).

Os valores dos ICC's da tabela acima, tanto para a avaliação das erosões dos VOI's quanto dos filtros nas imagens, foram utilizados para selecionar os atributos com maior reprodutibilidade pelo método não supervisionado, onde os atributos com ICC > 0,90 para as duas avaliações foram selecionados (atributos de 01 a 03). Já os valores dos *p-value* da tabela acima foram utilizados para a seleção supervisionada dos atributos, em que atributos com *p-value* < 0,05 foram selecionados (atributos de 04 a 09).

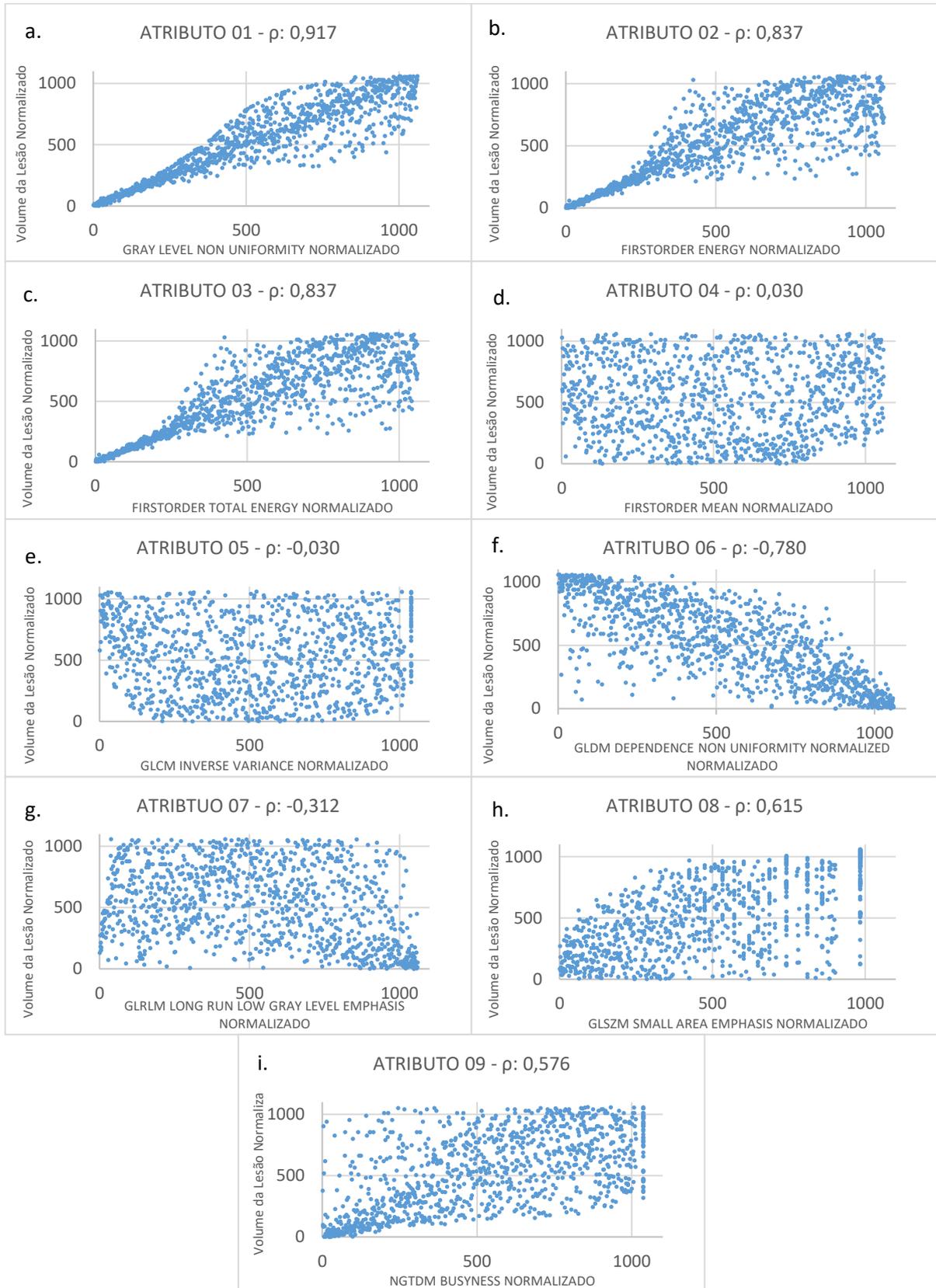
Outra característica levada em conta neste trabalho foi o volume da lesão. Entendendo que o volume da lesão tem por característica ser imutável (e, por este motivo, não pode ser harmonizado matematicamente), os atributos fortemente correlacionados ao volume podem não ser corretamente harmonizadas pelo método ComBat.

Para os pacientes do Hospital B, foram observadas lesões com volumes maiores (em relação aos pacientes dos hospitais A e C), além de maior dispersão entre os volumes, conforme é possível observar na Figura 17. Uma possível explicação para este achado é que o Hospital B, por ser um hospital de referência em doenças infecciosas da região na qual é localizado, recebe pacientes mais graves e consequentemente com lesões em estágios mais avançados.

Desta forma, os atributos considerados fortemente correlacionados com o volume da lesão não foram harmonizados pelo método ComBat, pois esta harmonização poderia gerar resultados inautênticos para tais atributos.

A Figura 18 apresenta os gráficos de dispersão entre os dados normalizados dos atributos selecionados e os volumes normalizados das lesões. As normalizações foram realizadas utilizando o cálculo do ranking médio (ranking de Besson) entre o atributo o volume da lesão.

Figura 18 – Gráficos de dispersões dos atributos normalizados do a. Atributo 01, b. Atributo 02, c. Atributo 03, d. Atributo 04, e. Atributo 05, f. Atributo 06, g. Atributo 07, h. Atributo 08 e i. Atributo 09 em função do volume da lesão normalizado



Fonte: O autor (2022).

Observando os gráficos de dispersão da Figura 18 e os valores das correlações de Spearman na Tabela 3 nota-se que os três atributos selecionados pelo método não supervisionado (atributo 01, 02 e 03) são fortemente correlacionados ao volume da lesão. Por este motivo, estes atributos não foram harmonizados pelo método ComBat. Embora a correlação seja mais forte para volumes menores das lesões (< 500 volumes normalizados pelo ranking de Besson) ainda se observa uma forte associação em volumes maiores.

Além disso, dos seis atributos selecionados pelo método supervisionado apresentados na Tabela 3, apenas o atributo 06 apresentou um valor de ρ que o correlaciona fortemente com o volume da lesão. Entretanto, observando o gráfico de dispersão do Atributo 06 em função do volume da lesão (Figura 21f) é possível notar que a correlação deste atributo com o volume da lesão ocorre com maior força em volumes menores das lesões. Mas, para volumes maiores das lesões (> 500 volumes normalizados pelo ranking de Besson) não é possível observar uma clara correlação, inclusive nos maiores volumes o valor do atributo é relativamente constante. Sendo assim, os seis atributos selecionados pelo método supervisionado foram harmonizados pelo método ComBat.

4.2 HARMONIZAÇÃO DOS ATRIBUTOS RADIÔMICOS

A seguir são apresentados os resultados dos atributos selecionados pelo método supervisionado e pelo método não supervisionado.

4.2.1 Análise dos atributos selecionados pelo método não supervisionado

O Atributo 01 (GLDM GRAY LEVEL NON UNIFORMITY) foi considerado fortemente correlacionado com o volume da lesão ($\rho = 0,917$), conforme apresentado no gráfico de dispersão da Figura 18a. Por este motivo o método ComBat de harmonização não foi aplicado para este atributo neste trabalho. Contudo, os seus valores médios e os desvios padrões sem a harmonização (sem ComBat) são apresentados na Tabela 4.

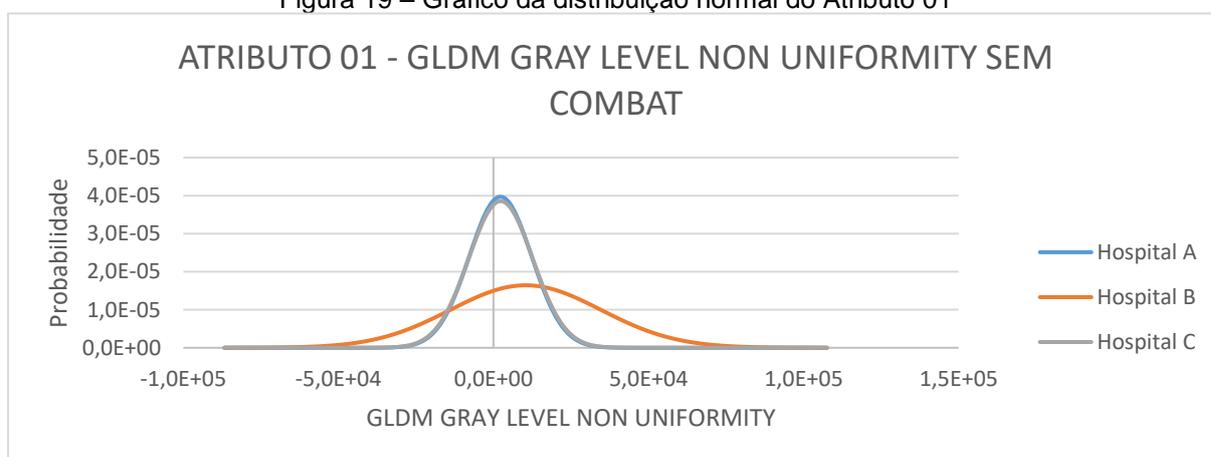
Tabela 4 – Dados do Atributo 01

ATRIBUTO 01 - GLDM GRAY LEVEL NON UNIFORMITY			
Sem ComBat			
Hospital	Hospital A	Hospital B	Hospital C
Média	$2,3 \times 10^3$	$1,0 \times 10^4$	$2,3E \times 10^3$
Desvio Padrão	$1,0 \times 10^4$	$2,4 \times 10^4$	$1,0 \times 10^4$

Fonte: O autor (2022).

Para melhor visualização dos resultados foi plotado um gráfico da distribuição normal do Atributo 01 sem a harmonização (sem ComBat) apresentado na Figura 19.

Figura 19 – Gráfico da distribuição normal do Atributo 01



GLDM: *Gray-Level Distance Matrix*. Fonte: O autor (2022).

O Atributo 01 (GLDM GRAY LEVEL NON UNIFORMITY), selecionado pelo método não supervisionado, é um atributo baseado em textura que evidencia a matriz de níveis de cinza não uniformes. Nota-se uma semelhança entre a distribuição dos volumes das lesões para os pacientes dos três hospitais (Figura 17) e as distribuições do Atributo 01 para os três hospitais. Esse efeito é explicado pelo fato de que o Atributo 01 está fortemente correlacionado com o volume da lesão.

O Atributo 02 (FIRSTORDER ENERGY) foi considerado fortemente correlacionado com o volume da lesão ($\rho = 0,837$), conforme apresentado no gráfico de dispersão da figura 18b, por este motivo o método ComBat de harmonização não foi aplicado neste atributo para este trabalho. Entretanto, os seus valores médios e os desvios padrões sem a harmonização (sem ComBat) são apresentados na Tabela 5.

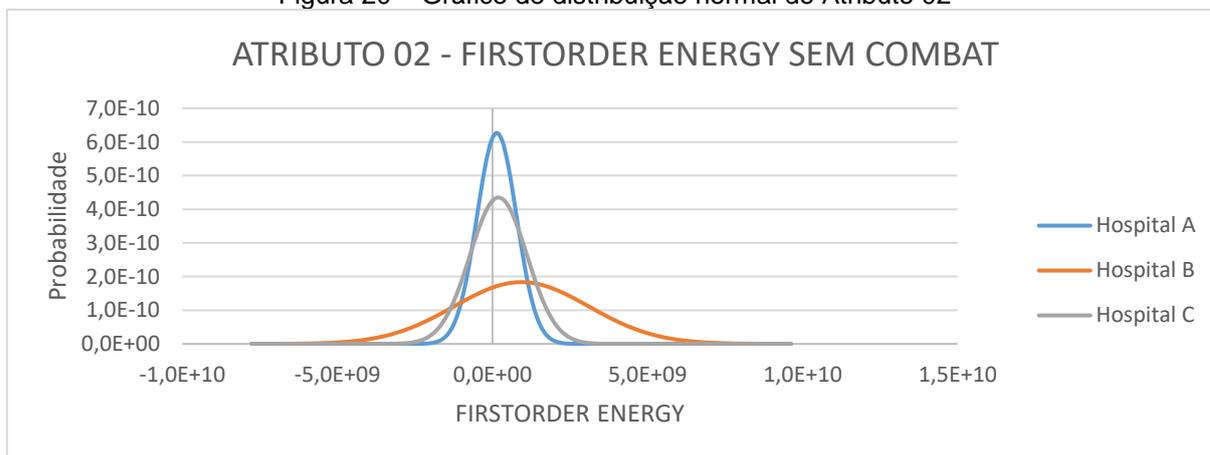
Tabela 5 – Dados do Atributo 02

ATRIBUTO 02 - FIRSTORDER ENERGY			
Sem ComBat			
Hospital	Hospital A	Hospital B	Hospital C
Média	$1,4 \times 10^8$	$9,3 \times 10^8$	$1,9 \times 10^8$
Desvio Padrão	$6,4 \times 10^8$	$2,2 \times 10^9$	$9,2 \times 10^8$

Fonte: O autor (2022).

Para melhor visualização dos resultados foi plotado um gráfico da distribuição normal do Atributo 02 sem a harmonização (sem ComBat) conforme exposto na Figura 20.

Figura 20 – Gráfico de distribuição normal do Atributo 02



Fonte: O autor (2022).

O Atributo 02 (FIRSTORDER ENERGY), selecionado pelo método não supervisionado, é um atributo de primeira ordem na qual sua análise é baseada no histograma de entropia dos níveis de cinza. Observa-se uma semelhança entre a distribuição dos volumes das lesões para os pacientes dos três hospitais (Figura 17) e a distribuições do Atributo 02 para os três hospitais. Esse efeito é explicado pelo fato de que o Atributo 01 está fortemente correlacionado com o volume da lesão.

O Atributo 03 (FIRSTORDER TOTAL ENERGY) foi considerado fortemente correlacionado com o volume da lesão ($\rho = 0,837$), conforme apresentado no gráfico de dispersão da Figura 18c, por este motivo o método ComBat de harmonização não foi aplicado neste atributo para este trabalho. Contudo, os seus valores médios e os desvios padrões sem a harmonização (sem ComBat) são apresentados na Tabela 6.

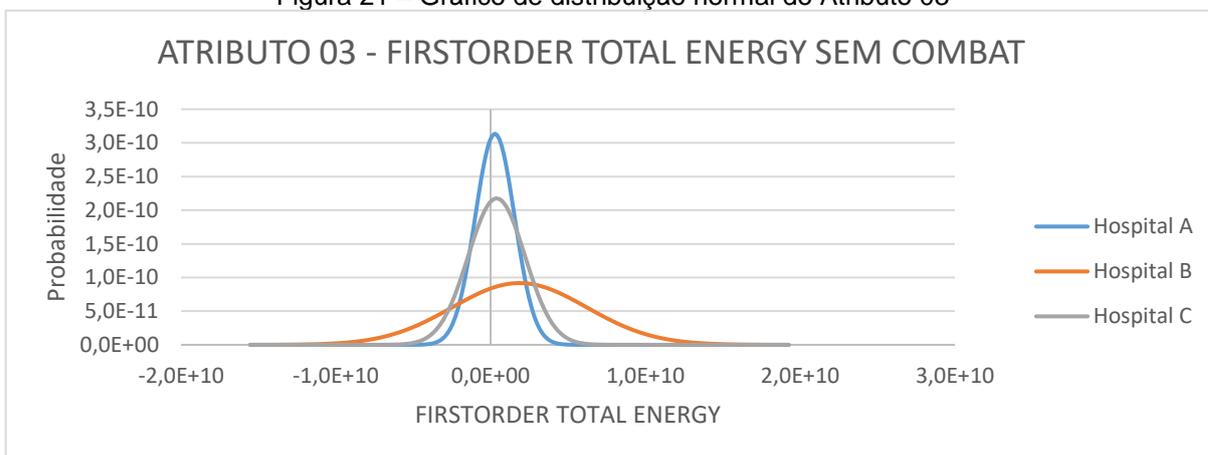
Tabela 6 – Dados do Atributo 03

ATRIBUTO 03 - FIRSTORDER TOTAL ENERGY			
Sem ComBat			
Hospital	Hospital A	Hospital B	Hospital C
Média	$2,8 \times 10^8$	$1,9 \times 10^9$	$3,9 \times 10^8$
Desvio Padrão	$1,3 \times 10^9$	$4,3 \times 10^9$	$1,8 \times 10^9$

Fonte: O autor (2022).

Para melhor visualização dos resultados, foi plotado um gráfico da distribuição normal do Atributo 03 sem a harmonização (sem ComBat) apresentado na Figura 21.

Figura 21 – Gráfico de distribuição normal do Atributo 03



Fonte: O autor (2022).

O Atributo 03 (FIRSTORDER TOTAL ENERGY), o último atributo selecionado pelo método não supervisionado, é um atributo de primeira ordem na qual sua análise é baseada no histograma de entropia total dos níveis de cinza. Além disso, ele está fortemente correlacionado com o Atributo 02, por este motivo os seus dados são semelhantes. Nota-se uma semelhança entre a distribuição dos volumes das lesões para os pacientes dos três hospitais (Figura 17) e a distribuições do Atributo 03 para os três hospitais. Esse efeito é explicado pelo fato de que o Atributo 01 está fortemente correlacionado com o volume da lesão.

O método de seleção não supervisionado avaliou a reprodutibilidade do atributo com relação a variações devido a diferentes segmentadores e protocolos. Observou-se que o método não supervisionado seleciona atributos harmonizados, sem necessidade de aplicação de harmonização numérica. Este fenômeno pode ser

explicado pela incorporação do método de borramento e melhoramento das imagens ao simular novos observadores para o cálculo do ICC: a implementação de um valor de corte alto ($ICC > 0,90$) implica na inclusão de atributos que sejam invariáveis com relação à qualidade das imagens.

Os atributos selecionados pelo método não supervisionado, a despeito da alta reprodutibilidade, demonstraram menor correlação clínica, conforme observado pelo *p-value* mais elevado.

4.2.2 Análise dos atributos selecionados pelo método supervisionado

O Atributo 04 (FIRSTORDER MEAN) foi considerado fracamente correlacionado com o volume da lesão ($\rho = 0,030$), conforme apresentado no gráfico de dispersão da Figura 18d, desta forma foi harmonizado para os três hospitais participantes da pesquisa. Os seus valores médios e os desvios padrões sem a harmonização (sem ComBat) e com a harmonização (com ComBat) são apresentados na Tabela 7.

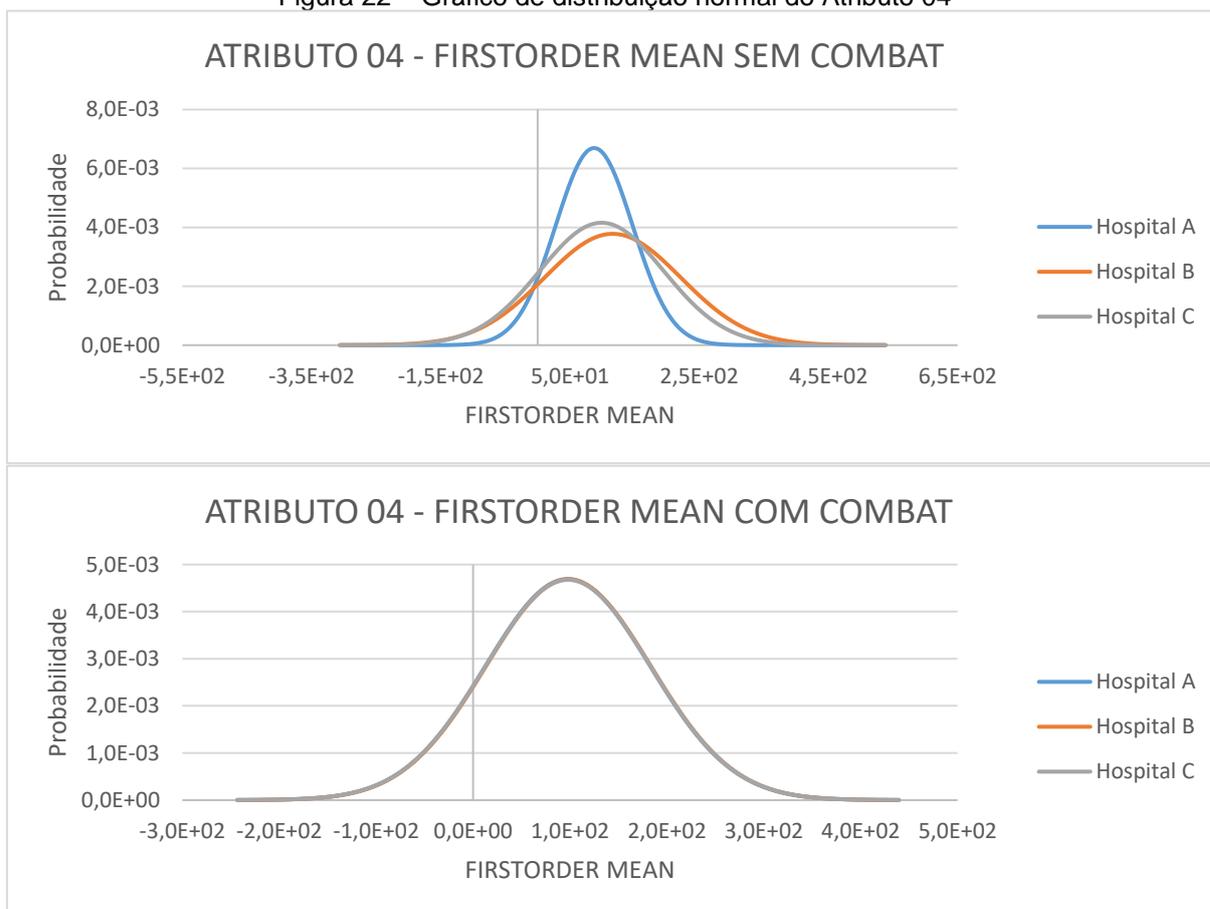
Tabela 7 – Dados do Atributo 04

ATRIBUTO 04 - FIRSTORDER MEAN						
	Sem ComBat			Com ComBat		
Hospital	Hospital A	Hospital B	Hospital C	Hospital A	Hospital B	Hospital C
Média	$8,7 \times 10^1$	$1,2 \times 10^2$	$9,9 \times 10^1$	$9,8 \times 10^1$	$9,8 \times 10^1$	$9,8 \times 10^1$
Desvio Padrão	$5,9 \times 10^1$	$1,1 \times 10^2$	$9,6 \times 10^1$	$8,5 \times 10^1$	$8,5 \times 10^1$	$8,5 \times 10^1$

Fonte: O autor (2022).

Para melhor visualização dos resultados, foram plotados os gráficos da distribuição normal do Atributo 04 sem a harmonização (sem ComBat) e com a harmonização (com ComBat) conforme exibido na Figura 22.

Figura 22 – Gráfico de distribuição normal do Atributo 04



Fonte: O autor (2022).

O Atributo 04 (FIRSTORDER MEAN), selecionado pelo método supervisionado, é um atributo de primeira ordem baseado no histograma médio de níveis de cinza. Sem a harmonização foi observado um coeficiente de variação de 11,7% entre os valores médios dos atributos dos três hospitais, após a harmonização ComBat o coeficiente de variação reduziu para 0,3%. É possível observar uma melhora significativa dos dados do Atributo 04 para os três hospitais após a harmonização ComBat e por este motivo pode ser utilizado clinicamente.

O Atributo 05 (GLCM INVERSE VARIANCE) foi considerado fracamente correlacionado com o volume da lesão ($\rho = -0,030$), conforme apresentado no gráfico de dispersão da Figura 18e, desta forma foi harmonizado para os três hospitais participantes da pesquisa. Os seus valores médios e os desvios padrões sem a harmonização (sem ComBat) e com a harmonização (com ComBat) são apresentados na Tabela 8.

Tabela 8 – Dados do Atributo 05

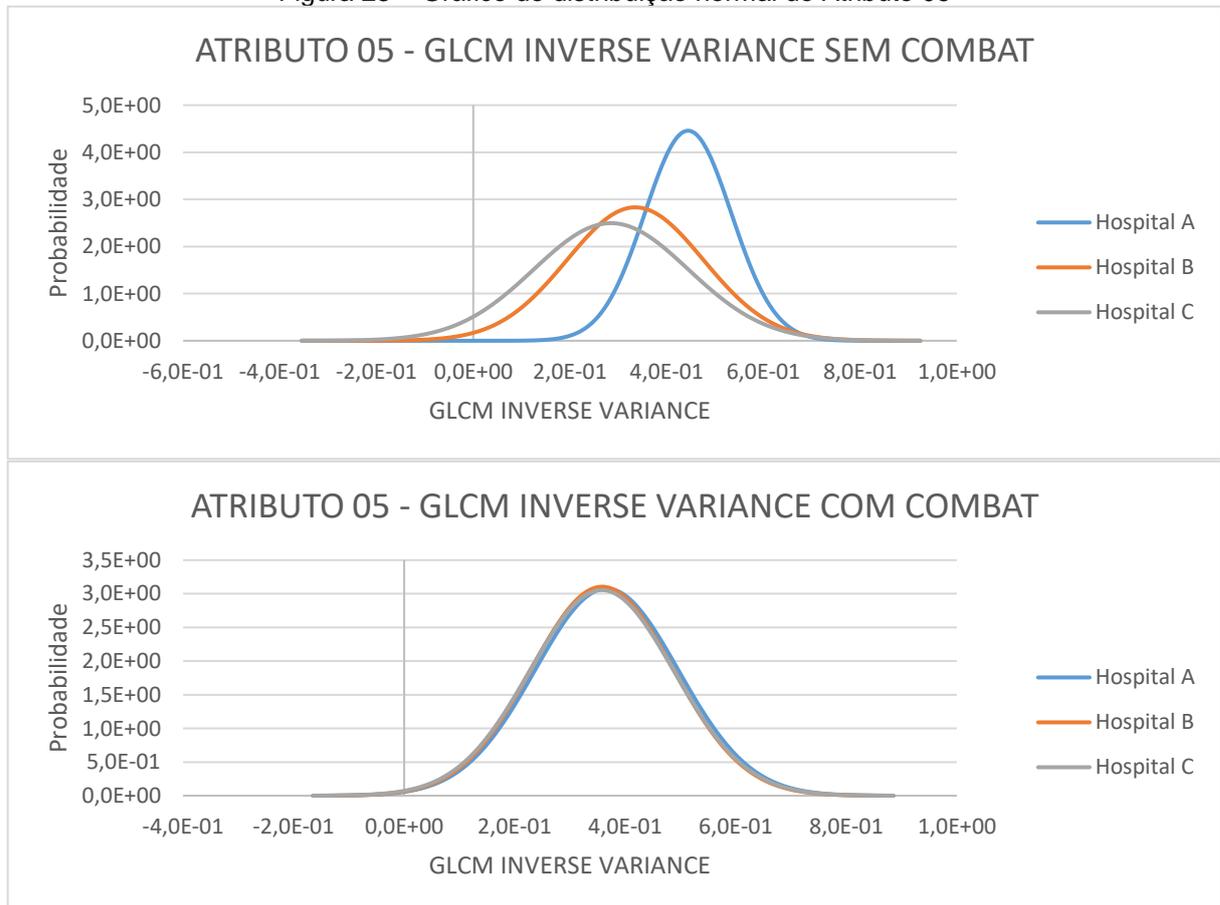
ATRIBUTO 05 - GLCM INVERSE VARIANCE

	Sem ComBat			Com ComBat		
Hospital	Hospital A	Hospital B	Hospital C	Hospital A	Hospital B	Hospital C
Média	$4,4 \times 10^{-1}$	$3,3 \times 10^{-1}$	$2,8 \times 10^{-1}$	$3,7 \times 10^{-1}$	$3,6 \times 10^{-1}$	$3,6 \times 10^{-1}$
Desvio Padrão	$8,9 \times 10^{-2}$	$1,4 \times 10^{-1}$	$1,6 \times 10^{-1}$	$1,3 \times 10^{-1}$	$1,3 \times 10^{-1}$	$1,3 \times 10^{-1}$

Fonte: O autor (2022).

Para melhor visualização dos resultados foram plotados os gráficos da distribuição normal do Atributo 05 sem a harmonização (sem ComBat) e com a harmonização (com ComBat) apresentado na Figura 23.

Figura 23 – Gráfico de distribuição normal do Atributo 05

GLCM: *Gray-Level Cooccurrence Matrix*. Fonte: O autor (2022).

O Atributo 05 (GLCM INVERSE VARIANCE), selecionado pelo método supervisionado, é um atributo baseado em textura na qual reflete uma matriz de coocorrência do inverso da variância do nível de cinza. Sem a harmonização, foi observado um coeficiente de variação de 18,8% entre os valores médios dos atributos

dos três hospitais, após a harmonização ComBat o coeficiente de variação reduziu para 1,1%. É possível observar uma melhora significativa no alinhamento dos dados do Atributo 05 para os três após a harmonização ComBat e por este motivo pode ser utilizado clinicamente.

O Atributo 06 (GLDM DEPENDENCE NON UNIFORMITY NORMALIZED) foi considerado fortemente correlacionado com o volume da lesão de acordo o coeficiente de correlação de Spearman ($\rho = -0,780$), entretanto, observando o gráfico de dispersão da Figura 18f nota-se que este atributo somente está fortemente correlacionado com o volume da lesão em pequenas áreas, sendo que em sua maioria, há uma elevada dispersão do atributo com o volume da lesão. Por este motivo, o Atributo 06 foi considerado fracamente correlacionado com o volume da lesão e harmonizado para os três hospitais participantes da pesquisa. Os seus valores médios e os desvios padrões sem a harmonização (sem ComBat) e com a harmonização (com ComBat) são apresentados na Tabela 9.

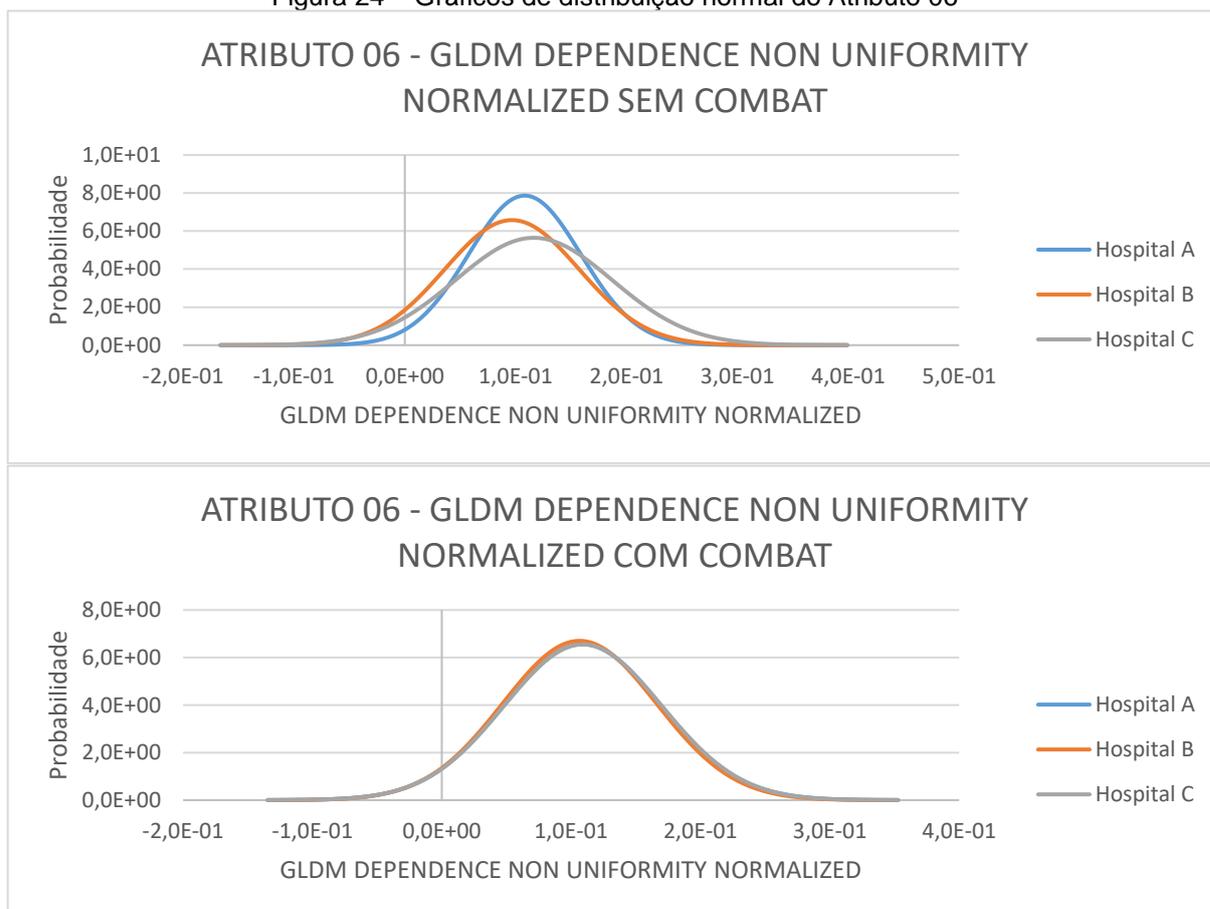
Tabela 9 – Dados do Atributo 06

ATRIBUTO 06 GLDM DEPENDENCE NON UNIFORMITY NORMALIZED						
	Sem ComBat			Com ComBat		
Hospital	Hospital A	Hospital B	Hospital C	Hospital A	Hospital B	Hospital C
Média	$1,1 \times 10^{-1}$	$9,7 \times 10^{-2}$	$1,2 \times 10^{-1}$	$1,1 \times 10^{-1}$	$1,1 \times 10^{-1}$	$1,1 \times 10^{-1}$
Desvio Padrão	$5,1 \times 10^{-2}$	$6,1 \times 10^{-2}$	$7,1 \times 10^{-2}$	$6,1 \times 10^{-2}$	$5,9 \times 10^{-2}$	$6,1 \times 10^{-2}$

Fonte: O autor (2022).

Para melhor visualização dos resultados, foram plotados os gráficos da distribuição normal do Atributo 06 sem a harmonização (sem ComBat) e com a harmonização (com ComBat) apresentado na Figura 24.

Figura 24 – Gráficos de distribuição normal do Atributo 06



GLDM: *Gray-Level Distance Matrix*. Fonte: O autor (2022).

O Atributo 06 (GLDM DEPENDENCE NON UNIFORMITY NORMALIZED), selecionado pelo método supervisionado, é um atributo baseado em textura na qual evidencia uma matriz normalizada da dependência de não uniformidade. Sem a harmonização foi observado um coeficiente de variação de 7,6% entre os valores médios dos atributos dos três hospitais, após a harmonização ComBat o coeficiente de variação reduziu para 1,0%. É possível observar a melhora no alinhamento dos dados do Atributo 06 para os três após a harmonização ComBat e por este motivo pode ser utilizado clinicamente.

O Atributo 07 (GLRLM LONG RUN LOW GRAY LEVEL EMPHASIS) foi considerado fracamente correlacionado com o volume da lesão ($\rho = -0,312$), conforme apresentado no gráfico de dispersão da Figura 18g, desta forma foi harmonizado para os três hospitais participantes da pesquisa. Os seus valores médios e os desvios padrões sem a harmonização (sem ComBat) e com a harmonização (com ComBat) são apresentados na Tabela 10.

Tabela 10 – Dados do Atributo 07

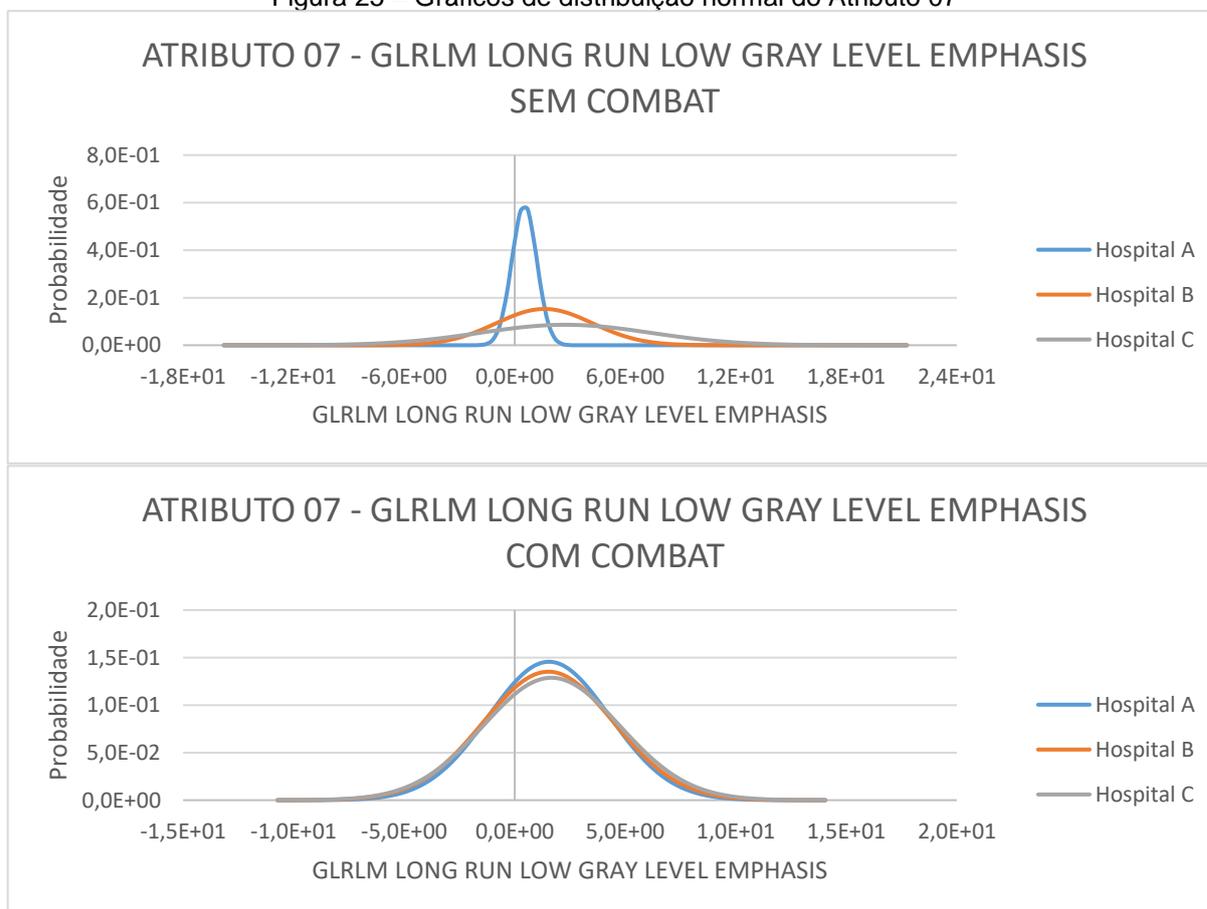
ATRIBUTO 07 - GLRLM LONG RUN LOW GRAY LEVEL EMPHASIS

	Sem ComBat			Com ComBat		
Hospital	Hospital A	Hospital B	Hospital C	Hospital A	Hospital B	Hospital C
Média	$5,2 \times 10^{-1}$	1,6	2,7	1,5	1,5	1,7
Desvio Padrão	$6,7 \times 10^{-1}$	2,6	4,6	2,7	2,9	3,1

Fonte: O autor (2022).

Para melhor visualização dos resultados, foram plotados os gráficos da distribuição normal do Atributo 07 sem a harmonização (sem ComBat) e com a harmonização (com ComBat) apresentado Figura 25.

Figura 25 – Gráficos de distribuição normal do Atributo 07

GLRLM: *Gray-Level Run-length Matrix*. Fonte: O autor (2022).

O Atributo 07 (GLRLM LONG RUN LOW GRAY LEVEL EMPHASIS), selecionado pelo método supervisionado, é um atributo baseado em textura na qual evidencia uma matriz de comprimento de execução de baixos níveis de cinza. Sem a harmonização foi observado um coeficiente de variação de 55,9% entre os valores médios dos

atributos dos três hospitais, após a harmonização ComBat o coeficiente de variação reduziu para 3,7%. É possível observar uma melhora significativa no alinhamento dos dados do Atributo 07 para os três hospitais após a harmonização ComBat por este motivo pode ser utilizado clinicamente.

O Atributo 08 (GLSZM SMALL AREA EMPHASIS) foi considerado fracamente correlacionado com o volume da lesão ($\rho = 0,615$), conforme apresentado no gráfico de dispersão da Figura 18h, desta forma foi harmonizado para os três hospitais participantes da pesquisa. Os seus valores médios e os desvios padrões sem a harmonização (sem ComBat) e com a harmonização (com ComBat) são apresentados na Tabela 11.

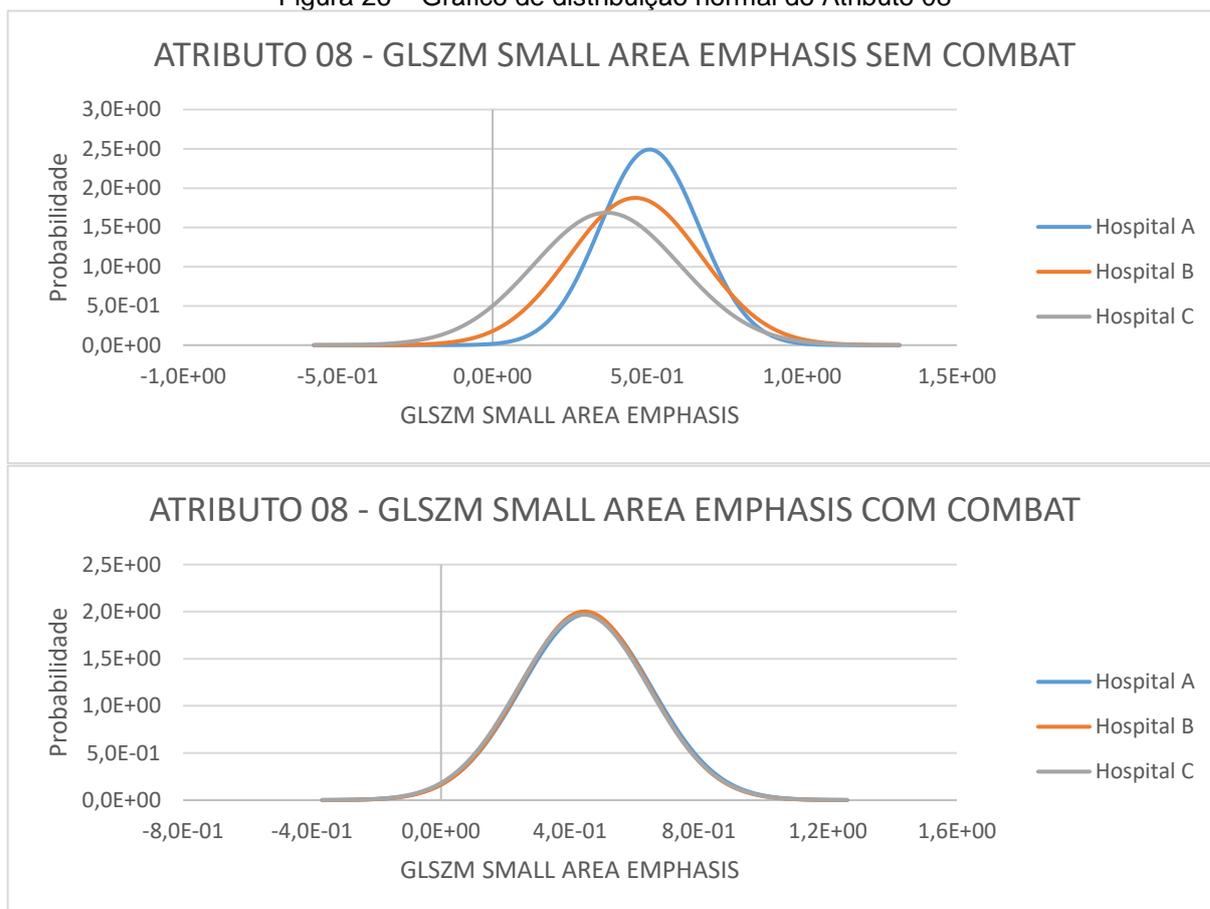
Tabela 11 – Dados do Atributo 08

ATRIBUTO 08 - GLSZM SMALL AREA EMPHASIS						
	Sem ComBat			Com ComBat		
Hospital	Hospital A	Hospital B	Hospital C	Hospital A	Hospital B	Hospital C
Média	$5,1 \times 10^{-1}$	$4,6 \times 10^{-1}$	$3,7 \times 10^{-1}$	$4,5 \times 10^{-1}$	$4,5 \times 10^{-1}$	$4,4 \times 10^{-1}$
Desvio Padrão	$1,6 \times 10^{-1}$	$2,1 \times 10^{-1}$	$2,4 \times 10^{-1}$	$2,0 \times 10^{-1}$	$1,9 \times 10^{-1}$	$2,0 \times 10^{-1}$

Fonte: O autor (2022).

Para melhor visualização dos resultados, foram plotados os gráficos da distribuição normal do Atributo 08 sem a harmonização (sem ComBat) e com a harmonização (com ComBat) conforme exposto na Figura 26.

Figura 26 – Gráfico de distribuição normal do Atributo 08



GLSZM: *Gray-Level Size Zone Matrix*. Fonte: O autor (2022).

O Atributo 08 (GLSZM SMALL AREA EMPHASIS), selecionado pelo método supervisionado, é um atributo baseado em textura na qual enfatiza matriz de pequenas zonas de níveis de cinza. Sem a harmonização foi observado um coeficiente de variação de 12,9% entre os valores médios dos atributos dos três hospitais, após a harmonização ComBat o coeficiente de variação reduziu para 0,8%. É possível observar uma melhora significativa no alinhamento dos dados do Atributo 08 para os três hospitais após a harmonização ComBat por este motivo ele pode ser utilizado clinicamente.

O Atributo 09 (NGTDM BUSYNESS) foi considerado fracamente correlacionado com o volume da lesão ($\rho = 0,576$), conforme apresentado no gráfico de dispersão da Figura 18i, desta forma foi harmonizado para os três hospitais participantes da pesquisa. Os seus valores médios e os desvios padrões sem a harmonização (sem ComBat) e com a harmonização (com ComBat) são apresentados na Tabela 12.

Tabela 12 – Dados do Atributo 09

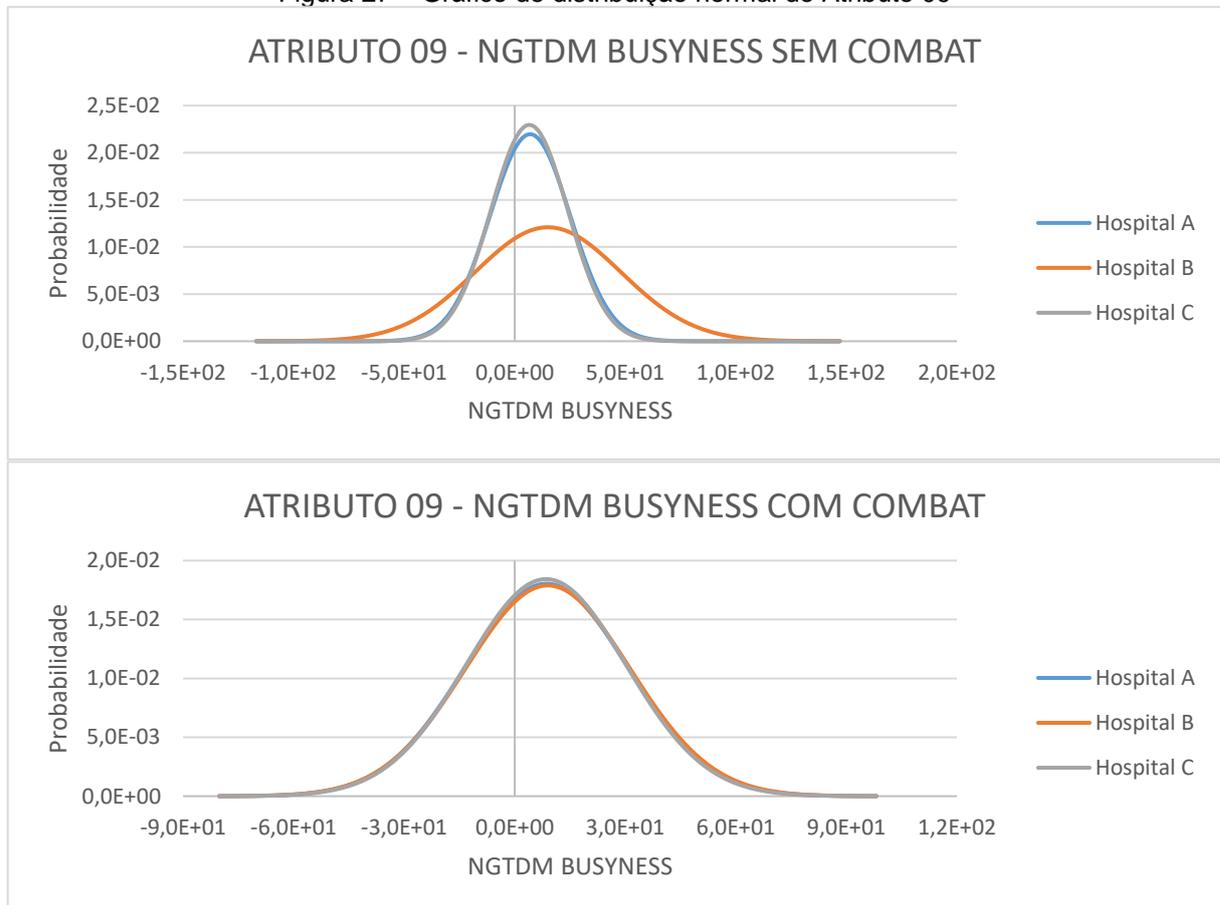
ATRIBUTO 09 - NGTDM BUSYNESS

	Sem ComBat			Com ComBat		
Hospital	Hospital A	Hospital B	Hospital C	Hospital A	Hospital B	Hospital C
Média	6,9	$1,5 \times 10^1$	6,6	8,6	8,9	8,5
Desvio Padrão	$1,8 \times 10^1$	$3,3 \times 10^1$	$1,7 \times 10^1$	$2,2 \times 10^1$	$2,2 \times 10^1$	$2,2 \times 10^1$

Fonte: O autor (2022).

Para melhor visualização dos resultados, foram plotados os gráficos da distribuição normal do Atributo 09 sem a harmonização (sem ComBat) e com a harmonização (com ComBat) apresentado na Figura 27.

Figura 27 – Gráfico de distribuição normal do Atributo 09

NGTDM: *Neighborhood Gray-Tone Difference Matrix*. Fonte: O autor (2022).

O Atributo 09 (NGTDM BUSYNESS), selecionado pelo método supervisionado, é um atributo baseado em textura na qual reflete mudanças rápidas de nível de cinza entre o pixel central ou voxel central e seus vizinhos. Sem a harmonização foi observado um coeficiente de variação de 40,8% entre os valores médios dos atributos

dos três hospitais, após a harmonização ComBat o coeficiente de variação reduziu para 2,2%. É possível observar que sem a harmonização ComBat os valores médios do Atributo 09 para os hospitais A e C estão alinhados, diferenciando apenas dos dados do hospital da B. Entretanto, após a harmonização ComBat foi observado uma melhora significativa dos dados do Atributo 09 para os três hospitais, por este motivo pode ser utilizado clinicamente.

É possível notar que os dados radiômicos dos atributos selecionados pelo método supervisionado sem a harmonização ComBat dos três hospitais são distintos, com exceção do Atributo 09 em que os dados radiômicos dos hospitais A e C são semelhantes. Além disso, todos os atributos selecionados pelo método supervisionado foram harmonizados com o método ComBat. Observou-se um significativo alinhamento dos dados radiômicos dos atributos harmonizados pelo método ComBat, sendo estes fortemente elegíveis para o uso clínico.

É importante destacar que o método de seleção supervisionado está associado a um conjunto de lesões ao status do paciente (COVID-19 / não COVID-19) e não é possível identificar o VOI que de fato tem características de COVID-19, pois um mesmo paciente pode ter lesões com características de COVID-19 ou não. Neste caso, pressupõe-se que é mais provável que um mesmo paciente tenha mais chances de que a maioria dos VOI's tenham as características de sua doença. Assim, o método supervisionado utilizado neste trabalho demonstra uma associação do atributo com a doença, mas que não é possível estimar a intensidade desta associação, visto que esta análise é no nível da lesão (e não no nível do paciente).

Observando os resultados dos dados dos atributos radiômicos que foram harmonizados pelo método ComBat nota-se a eficácia desta harmonização. Segundo Orlhac e colaboradores (2019), o método ComBat de harmonização é um procedimento promissor para viabilização de estudos radiômicos realizados em diferentes centros de radiologia.

Orlhac e colaboradores (2019) utilizaram simuladores de imagens e exames tomográficos de pacientes reconstruídos com diferentes protocolos de TC para avaliar o método ComBat de harmonização. Foi demonstrado que o ComBat removeu o efeito provocado pelos diferentes equipamentos e dos diferentes protocolos de aquisição, preservando as diferenças entre os padrões de textura. A correção do efeito dos diferentes equipamentos foi confirmada usando imagens de pacientes reconstruídas com diferentes protocolos de imagem.

Além disso, Fortin e colaboradores (2018) investigaram os efeitos da variabilidade das medidas de espessura cortical em imagens obtidas em diferentes equipamentos de ressonância magnética, utilizando dados de dois grandes estudos multicêntricos com ressonância magnética (*EMBARC study* com 4 equipamentos e *VDLC study* com 7 equipamentos). O método ComBat foi utilizado para harmonizar valores de espessura cortical entre os 11 equipamentos. Foi demonstrado que o ComBat é eficaz na remoção da variabilidade associada aos diferentes equipamentos, preservando a intercomparação entre as medidas da espessura cortical dos participantes.

Outros métodos de harmonização podem ser aplicados, por exemplo utilizando-se pós processamento de imagens. Neste caso, a harmonização numérica é substituída por uma transformação das imagens, tal que os atributos entre os equipamentos sejam comparáveis (MACHADO et al., 2019). Neste contexto, os atributos encontrados pelo método supervisionado, por sua provável importância clínica, podem ser harmonizados através deste método. Desta forma, quando o ComBat não estiver disponível em ensaios futuros, o método de pós processamento poderá ser avaliado ajustando-se os filtros na reconstrução das imagens até que as imagens obtidas por diferentes equipamentos resultem em atributos no mesmo domínio.

5 CONCLUSÃO

Para os atributos radiômicos selecionados pelo método não supervisionado, que enfatiza os atributos mais reprodutíveis, observou-se que estes são fortemente correlacionados com o volume da lesão. Por este motivo, para este trabalho nenhum dos atributos selecionados por este método foi harmonizado utilizando o ComBat. Entretanto, notou-se que esses atributos possuem seus dados alinhados para os Hospitais A e C sem a harmonização ComBat, diferenciando apenas dos dados do Hospital B, devido a grande diferença de volume da lesão dos pacientes deste hospital. Conclui-se que o método de seleção não supervisionado demonstrou ser eficiente na triagem de atributos reprodutíveis entre os hospitais, não sendo necessário a realização da harmonização ComBat.

Já para os atributos de diferentes classes selecionados pelo método supervisionado, que enfatiza os atributos com melhores predição a doença, foi possível observar que após a harmonização ComBat houve uma melhora significativa na distribuição dos dados dos atributos radiômicos para os três hospitais, para os atributos fracamente relacionados com o volume da lesão.

Neste trabalho, demonstramos as características dos dois métodos de seleção propostos, considerando suas peculiaridades na robustez dos atributos quanto à reprodutibilidade, harmonização, correlação clínica e viés populacional. Nove atributos radiômicos foram selecionados para aplicações multicêntricas em COVID-19, sendo 03 atributos diretamente aplicáveis, e 06 atributos aplicáveis após harmonização pelo método ComBat.

O presente trabalho possui um viés que foi considerado nas análises: as lesões dos participantes do Hospital B possuem um maior volume em relação às lesões dos participantes dos Hospitais A e C. Sendo assim, a correlação entre o atributo e o volume da lesão foi analisada criticamente para a aplicação da harmonização ComBat.

REFERÊNCIAS

- BRASIL. Diagnóstico e Tratamento da COVID-19. **Ministério da Saúde**. 2020. v.1, p. 1–398.
- BRASIL. Diretrizes Brasileiras para Tratamento Hospitalar do Paciente com COVID-19. **Ministério da Saúde**. [S.l.]: [s.n.], 2021, p. 1–53.
- BUSHBERG, J. T. **The Essential Physics of Medical Imaging**. 2012. V. 3.
- BUSHBERG, J. T.. **The Essential Physics of Medical Imaging**. 2021. V. 4.
- BUSHONG, S. C. **Radiologic Science for Technologists: Physics, Biology and Protection**. [S.l.]: [s.n.], 2008.
- FANG, M.; HE, B.; LI, L.; DONG, D.; YANG, X.; LI, C.; MENG, L.; ZHONG, L.; LI, H.; TIAN, J. CT radiomics can help screen the Coronavirus disease 2019 (COVID-19): a preliminary study. **Science China Information Sciences**, 2020. v. 63, n. 7, p. 1–8.
- FORTIN, J. P. *et al.* Harmonization of cortical thickness measurements across scanners and sites. **NeuroImage**, 2018. v. 167, n. June 2017, p. 104–120.
Disponível em: <<https://doi.org/10.1016/j.neuroimage.2017.11.024>>.
- GRIETHUYSEN, J. J. M. VAN *et al.* Computational radiomics system to decode the radiographic phenotype. **Cancer Research**, 2017. v. 77, n. 21, p. e104–e107.
- JOHNSON, W. E.; LI, Cheng; RABINOVIC, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. **Biostatistics**, 2007. v.8, n 1, p.118–127.
- KUMAR, V.; GU, Y.; BASU, S.; BERGLUND, A.; ESCHRICH, S.; SCHABATH, M.; FORSTER, K.; AERTS, H.; DEKKER, A.; FENSTERMACHER, D.; GOLDGOF, D.; HALL, L.; LAMBIN, P.; BALAGURUNATHAN, Y.; GATENBY, R.; GILLIES, R. Radiomics: The process and the challenges. **Magnetic Resonance Imaging**, 2012. v. 30, n. 9, p. 1234–1248.
- LAMBIN, P.; VELAZQUEZ, E.; LEIJENAAR, R.; CARVALHO, S.; STIPHOUT, R.; GRANTON, P.; ZEGERS, C.; GILLIES, R.; BOELLARD, R.; DEKKER, A.; AERTS, H. Radiomics: Extracting more information from medical images using advanced feature analysis. **European Journal of Cancer**, 2012. v. 48, n. 4, p. 441–446.
- MACHADO, M. A. D.; MENEZES, V. O.; NAMÍAS, M.; VIEIRA, N. S.; QUEIROZ, C. C.; MATHEOD, R.; ALESSIO, A. M.; OLIVEIRA, M. L. Protocols for harmonized quantification and noise reduction in low-dose oncologic 18F-FDG PET/CT imaging. **Journal of Nuclear Medicine Technology**, 2019. v. 47, n. 1, p. 47–54.
- MACHADO, M. A. D.; PEREIRA PITA, J. L.; MARTINS NETTO, E. PET Radiomic Features Variability: A Phantom Study on the Influence of Reconstruction and Discretization Method. **Revista Brasileira de Física Médica**, 2021. v. 15, p. 598.
- MAYERHOEFER, M. E.; MATERKA, A.; LANGS, G.; HAGGSTROM, I.;

SZCZYPINSKI, P.; GIBBS, P.; COOK, G. Introduction to radiomics. **Journal of Nuclear Medicine**, 2020. v. 61, n. 4, p. 488–495.

MEIRELLES, G. De S. P. COVID-19 : uma breve atualização para radiologistas. **Radiol Bras**, 2020. v. 53, n. 5, p. 320–28. Disponível em: <<http://dx.doi.org/10.1590/0100-3984.2020.0074>>.

MUKAKA, M. M. 81576-194640-1-Pb. **Malawi Medical Journal**, 2012. v. 24, n. September, p. 69–71. Disponível em: <<https://www.ajol.info/index.php/mmj/article/view/81576>>.

ORLHAC, F.; BOUGHDAD, S.; PHILIPPE, C.; BOURDILLON, H.; NIOCHE, C.; CHAMPION, L.; SOUSSAN, M.; FROUIN, F.; FROUIN, V.; BUVAT, I. A postreconstruction harmonization method for multicenter radiomic studies in PET. **Journal of Nuclear Medicine**, 2018. v. 59, n. 8, p. 1321–1328.

ORLHAC, F.; FROUIN, F.; NIOCHE, C.; AYACHE, N.; BUVAT, I. Validation of a method to compensate multicenter effects affecting CT radiomics. **Radiology**, 2019. v. 291, n. 1, p. 53–59.

PUGLIA, E. B. M. DI *et al.* Chest computed tomography in covid-19 pneumonia: A retrospective study of 155 patients at a university hospital in Rio de Janeiro, Brazil. **Radiologia Brasileira**, 2021. v. 54, n. 1, p. 9–14.

SANTOS, L. P. *et al.* Testagem Rápida para COVID-19 em uma rede de farmácias de uma cidade do Sudoeste da Bahia Rapid Testing for COVID-19 in a pharmacy chain in a city in the Southwest of Bahia Pruebas rápidas de COVID-19 en una cadena de farmacias en una ciudad del suroeste . 2021. v. 2021, p. 1–11.

SANTOS, M. K. *et al.* Inteligência artificial , aprendizado de máquina , diagnóstico auxiliado por computador e radiômica : avanços da imagem rumo à medicina de precisão. **Radiol Bras**, 2019. v. 52, n. 6, p. 387–396.

MOGAMI, R.; LOPES, A. J.; ARAÚJO, F. R. C.; ALMEIDA, F. C. S.; MESSEDER, A. M. C.; KOIFMAN, A. C. B.; GUIMARÃES, A. B.; MONTEIRO, A. Chest computed tomography in covid-19 pneumonia: A retrospective study of 155 patients at a university hospital in Rio de Janeiro, Brazil. **Radiologia Brasileira**, 2021. v. 54, n. 1, p. 9–14.

KOENIGKAM, S. M.; FERREIRA, JR.; WADA, D. T.; TENÓRIO, A. P. M.; NOGUEIRA, B. M. H.; AZEVEDO, M. P. M. A. Inteligência artificial, aprendizado de máquina, diagnóstico auxiliado por computador e radiômica: avanços da imagem rumo à medicina de precisão. **Radiol Bras**, 2019. v. 52, n. 6, p. 387–396.

WU, Q. *et al.* Radiomics analysis of computed tomography helps predict poor prognostic outcome in COVID-19. **Theranostics**, 2020. v. 10, n. 16, p. 7231–7244.

ZU, Z. Y.; JIANG, M.; XU, P. P.; CHEN, W.; NI, Q. Q.; LU, G. M.; ZHANG, L. J. Coronavirus Disease 2019 (COVID-19): A Perspective from China. **Radiology**, 2020. v. 296, n. 2, p. E15–E25.

APÊNDICE A – CÓDIGO PYTHON DE HARMONIZAÇÃO DOS ATRIBUTOS RADIÔMICOS

```

# -*- coding: utf-8 -*-
"""
Created on Wed Jan 27 15:54:26 2021

@author: mnamias
"""
from neuroComBat import neuroComBat
import pandas as pd
import numpy as np

csvdir = "C:\\Users\\Danillo Menezes\\Desktop\\atributos radiômicos\\"
csvfile = 'building.csv'
df = pd.read_csv(csvdir + csvfile)

df = df.replace(to_replace='UFCG', value=0)
df = df.replace(to_replace='UFJF', value=1)
df = df.replace(to_replace='UFBA', value=2)
df = df.replace(to_replace='HUPES', value=3)

features = np.array(df.iloc[:, -94:])
features = features.transpose()

covars = pd.DataFrame(df.iloc[:, 0:2])

batch_col = 'Hospital'
ref_batch = 2

# Arguments
# -----
# dat : a pandas data frame or numpy array
#       neuroimaging data to correct with shape = (features, samples)
#       e.g. cortical thickness measurements, image voxels, etc
# covars : a pandas data frame w/ shape = (samples, covariates)
#          demographic/phenotypic/behavioral/batch data

# batch_col : string indicating batch (scanner) column name in covars
# categorical_cols : string or list of strings of categorical variables
to adjust for
# - e.g. male or female
# continuous_cols : string or list of strings of continuous variables
to adjust for
# - e.g. age
# eb : should Empirical Bayes be performed?
# - True by default
# parametric : should parametric adjustments be performed?
# - True by default
# mean_only : should only be the mean adjusted (no scaling)?
# - False by default
# ref_batch : batch (site or scanner) to be used as reference for batch
adjustment.
# - False by default

# Returns
# -----
# - A numpy array with the same shape as `dat` which has now been
ComBat-harmonized

```

```
data_ComBat = neuroComBat(dat=features, covars=covars, batch_col=batch_col,  
ref_batch=ref_batch)  
  
data_ComBat = data_ComBat['data'].transpose()  
  
df.iloc[:, -94:] = data_ComBat  
  
csvfile_ComBat = 'features_ComBat_18_11_2021.csv'  
  
df.to_csv(csvdir + csvfile_ComBat)
```

APÊNDICE B – LISTA DOS ATRIBUTOS RADIÔMICOS

NOME DO ATRIBUTO	ICC - EROSÃO	ICC - FILTRO	P- VALUE
FIRSTORDER 10PERCENTILE	0,511	0,987	0,000
FIRSTORDER 90PERCENTILE	0,446	0,987	0,000
FIRSTORDER ENERGY	0,907	0,999	0,739
FIRSTORDER ENTROPY	0,583	0,825	0,089
FIRSTORDER INTERQUARTILERANGE	0,494	0,960	0,104
FIRSTORDER KURTOSIS	0,746	0,915	0,170
FIRSTORDER MAXIMUM	0,723	0,930	0,001
FIRSTORDER MEAN	0,502	0,998	0,000
FIRSTORDER MEANABSOLUTEDEVIATION	0,473	0,966	0,137
FIRSTORDER MEDIAN	0,519	0,996	0,000
FIRSTORDER MINIMUM	0,551	0,926	0,000
FIRSTORDER RANGE	0,738	0,904	0,039
FIRSTORDER ROBUSTMEANABSOLUTEDEVIATION	0,474	0,969	0,134
FIRSTORDER ROOTMEANSQUARED	0,533	0,989	0,000
FIRSTORDER SKEWNESS	0,622	0,854	0,289
FIRSTORDER TOTALENERGY	0,907	0,999	0,739
FIRSTORDER UNIFORMITY	0,539	0,643	0,114
FIRSTORDER VARIANCE	0,395	0,951	0,144
GLCM AUTOCORRELATION	0,733	0,774	0,026
GLCM CLUSTERPROMINENCE	0,360	0,898	0,122
GLCM CLUSTERSHADE	0,600	0,905	0,008
GLCM CLUSTERTENDENCY	0,440	0,964	0,146
GLCM CONTRAST	0,201	0,742	0,357
GLCM CORRELATION	0,523	0,429	0,470
GLCM DIFFERENCEAVERAGE	0,308	0,714	0,401
GLCM DIFFERENCEENTROPY	0,499	0,576	0,140
GLCM DIFFERENCEVARIANCE	0,280	0,671	0,139
GLCM ID	0,468	0,522	0,325
GLCM IDM	0,451	0,570	0,361
GLCM IDMN	0,384	0,431	0,443
GLCM IDN	0,477	0,370	0,188
GLCM IMC1	0,368	0,671	0,413
GLCM IMC2	0,503	0,811	0,224
GLCM INVERSEVARIANCE	0,582	0,150	0,000
GLCM JOINTAVERAGE	0,708	0,903	0,020
GLCM JOINTENERGY	0,538	0,436	0,065

GLCM JOINTENTROPY	0,628	0,743	0,058
GLCM MAXIMUMPROBABILITY	0,569	0,487	0,067
GLCM MCC	0,506	0,463	0,691
GLCM SUMAVERAGE	0,708	0,903	0,020
GLCM SUMENTROPY	0,607	0,802	0,037
GLCM SUMSQUARES	0,408	0,950	0,111
GLDM DEPENDENCEENTROPY	0,553	0,870	0,015
GLDM DEPENDENCENONUNIFORMITY	0,883	0,484	0,414
GLDM DEPENDENCENONUNIFORMITYNORMALIZED	0,151	0,480	0,000
GLDM DEPENDENCEVARIANCE	0,535	0,764	0,015
GLDM GRAYLEVELNONUNIFORMITY	0,903	0,951	0,098
GLDM GRAYLEVELVARIANCE	0,394	0,950	0,135
GLDM HIGHGRAYLEVEL EMPHASIS	0,720	0,773	0,027
GLDM LARGEDEPENDENCE EMPHASIS	0,624	0,531	0,529
GLDM LARGEDEPENDENCEHIGHGRAYLEVEL EMPHASIS	0,846	0,886	0,004
GLDM LARGEDEPENDENCELOWGRAYLEVEL EMPHASIS	0,502	0,244	0,020
GLDM LOWGRAYLEVEL EMPHASIS	0,537	0,476	0,004
GLDM SMALLDEPENDENCE EMPHASIS	0,213	0,608	0,744
GLDM SMALLDEPENDENCEHIGHGRAYLEVEL EMPHASIS	0,129	0,573	0,059
GLDM SMALLDEPENDENCELOWGRAYLEVEL EMPHASIS	0,349	0,416	0,593
GLRLM GRAYLEVELNONUNIFORMITY	0,897	0,889	0,075
GLRLM GRAYLEVELNONUNIFORMITYNORMALIZED	0,484	0,631	0,091
GLRLM GRAYLEVELVARIANCE	0,402	0,948	0,152
GLRLM HIGHGRAYLEVEL RUN EMPHASIS	0,708	0,772	0,032
GLRLM LONGRUN EMPHASIS	0,788	0,143	0,645
GLRLM LONGRUNHIGHGRAYLEVEL EMPHASIS	0,894	0,238	0,219
GLRLM LONGRUNLOWGRAYLEVEL EMPHASIS	0,516	0,075	0,001
GLRLM LOWGRAYLEVEL RUN EMPHASIS	0,518	0,515	0,005
GLRLM RUN ENTROPY	0,600	0,914	0,040
GLRLM RUNLENGTHNONUNIFORMITY	0,865	0,812	0,029
GLRLM RUNLENGTHNONUNIFORMITYNORMALIZED	0,442	0,661	0,196
GLRLM RUNPERCENTAGE	0,554	0,555	0,755
GLRLM RUNVARIANCE	0,806	0,157	0,310
GLRLM SHORTRUN EMPHASIS	0,525	0,366	0,539
GLRLM SHORTRUNHIGHGRAYLEVEL EMPHASIS	0,652	0,695	0,021
GLRLM SHORTRUNLOWGRAYLEVEL EMPHASIS	0,474	0,713	0,057

GLSZM GRAYLEVELNONUNIFORMITY	0,883	0,544	0,000
GLSZM GRAYLEVELNONUNIFORMITYNORMALIZED	0,482	0,596	0,360
GLSZM GRAYLEVELVARIANCE	0,547	0,921	0,098
GLSZM HIGHGRAYLEVELZONEEMPHASIS	0,706	0,783	0,025
GLSZM LARGEAREAEMPHASIS	0,927	0,141	0,145
GLSZM LARGEAREAHIGHGRAYLEVELEMPHASIS	0,931	0,203	0,089
GLSZM LARGEAREALOWGRAYLEVELEMPHASIS	0,411	0,020	0,642
GLSZM LOWGRAYLEVELZONEEMPHASIS	0,495	0,681	0,021
GLSZM SIZEZONENONUNIFORMITY	0,862	0,555	0,022
GLSZM SIZEZONENONUNIFORMITYNORMALIZED	0,223	0,263	0,395
GLSZM SMALLAREAEMPHASIS	0,417	0,262	0,013
GLSZM SMALLAREAHIGHGRAYLEVELEMPHASIS	0,633	0,670	0,027
GLSZM SMALLAREALOWGRAYLEVELEMPHASIS	0,437	0,193	0,322
GLSZM ZONEENTROPY	0,633	0,850	0,041
GLSZM ZONEPERCENTAGE	0,211	0,720	0,648
GLSZM ZONEVARIANCE	0,927	0,142	0,017
NGTDM BUSYNESS	0,620	0,072	0,000
NGTDM COARSENESS	0,142	0,127	0,000
NGTDM COMPLEXITY	0,812	0,548	0,023
NGTDM CONTRAST	0,129	0,702	0,734
NGTDM STRENGTH	0,046	0,894	0,061
SHAPE VOXELVOLUME	0,880	1,000	0,123

ANEXO A – PARECER CONSUBSTANCIADO DO CEP

<p>UFBA - HOSPITAL UNIVERSITÁRIO PROF. EDGARD SANTOS DA UNIVERSIDADE FEDERAL DA BAHIA & HUPES/UFBA</p>	
--	---

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Desenvolvimento e validação de software para diagnóstico de covid-19 baseado em Tomografia Computadorizada, Radiômica e Inteligência Artificial.

Pesquisador: Eduardo Martins Netto

Área Temática:

Versão: 1

CAAE: 36624620.0.0000.0049

Instituição Proponente: Hospital Universitário Prof. Edgard Santos-UFBA

Patrocinador Principal: FINANCIADORA DE ESTUDOS E PROJETOS - FINEP

DADOS DO PARECER

Número do Parecer: 4.231.669

Apresentação do Projeto:

Trata-se de um protocolo para o desenvolvimento de um software RIS/PACS com aplicação de radiômica e inteligência artificial que tenha capacidade de auxiliar no diagnóstico de Covid-19 em relação a outras doenças em pacientes com sintomas gripais.

Projeto multicêntrico, do tipo coorte retrospectiva, com 10 pacientes de cada instituição com diagnósticos relacionados à pneumonia anterior a outubro-2019, e mais 10 pacientes de cada instituição com diagnósticos relacionados à Covid-19 entre abril-2020 e junho-2020. Total de pacientes 80 de participantes retrospectivo.

Apresenta como desfecho primário uma inovação incremental ao processo de diagnóstico por imagens TC, induzida por uma forte inovação de produtos RIS/PACS ao incorporar técnicas de radiômica e inteligência artificial.

O advento da radiômica e da inteligência artificial representa uma quebra de paradigma na medicina, pois tem possibilitado melhorias importantes nos modelos diagnósticos de várias doenças, assim como vem contribuindo com a abordagem de medicina personalizada.

Embora alguns padrões da TC sejam mais comuns em pacientes com Covid-19 em relação à outras infecções, padrões inespecíficos representam uma limitação da TC. O emergente mercado da radiômica e inteligência artificial como ferramentas para apoio ao diagnóstico pode ser uma

Endereço: Rua Augusto Viana, s/nº - 1º Andar		CEP: 40.110-060
Bairro: Canela	Município: SALVADOR	
UF: BA	Telefone: (71)3283-8043	Fax: (71)3283-8140
		E-mail: cep.hupes@ebserh.gov.br

UFBA - HOSPITAL
UNIVERSITÁRIO PROF.
EDGARD SANTOS DA
UNIVERSIDADE FEDERAL DA
BAHIA & HUPES/UFBA



Continuação do Parecer: 4.231.669

forte aliada no combate à Covid-19, pois desvendam informações relevantes sobre a doença, imperceptíveis ao olho humano. Este projeto visa o desenvolvimento de um software para quantificação de imagens e inteligência artificial para o diagnóstico do Covid-19, embarcado em um sistema RIS/PACS na nuvem.

Objetivo da Pesquisa:

Objetivo Primário:

Otimizar o diagnóstico do Covid-19 por tomografia computadorizada a partir de melhoramento tecnológico de um software RIS/PACS com aplicação de radiômica e inteligência artificial.

Objetivo Secundário:

- Desenvolver um aplicativo para segmentação automática de regiões de interesse (lesões de Covid-19) de imagens de tomografia computadorizada, embarcado em um sistema PACS (Picture Archiving and Communication System);
- Desenvolver um software para realizar a quantificação radiômica (padrões de imagens) de lesões pulmonares do Covid-19;
- Desenvolver painel de mineração de dados clínicos (achados dos laudos e fatores de risco) para aplicação em modelos preditivos;
- Desenvolver escores de risco usando dados clínicos, padrões de radiômica da tomografia computadorizada e inteligência artificial (IA).

Avaliação dos Riscos e Benefícios:

Riscos:

Os riscos associados com o armazenamento das imagens pode ser a perda de confidencialidade, porém elas estarão sob sigilo e com código sob acesso somente dos pesquisadores. O desenvolvimento do trabalho implica em riscos mínimos de perda de confidencialidade para os participantes da pesquisa. Estes riscos serão minimizados utilizando-se códigos de acesso nos computadores. Os dados clínicos serão tratados como sigilosos.

Os questionários serão codificados. Uma lista mestra ligará os nomes e identificação dos pacientes aos questionários de pesquisa (que usarão códigos). Portanto, a identidade dos pacientes será preservada. Todo o cuidado será tomado quanto ao sigilo dos dados (senhas de acesso: ao

Endereço: Rua Augusto Viana, s/nº - 1º Andar			
Bairro: Canela		CEP: 40.110-060	
UF: BA	Município: SALVADOR		
Telefone: (71)3283-8043	Fax: (71)3283-8140	E-mail: cep.hupes@ebserh.gov.br	

UFBA - HOSPITAL
UNIVERSITÁRIO PROF.
EDGARD SANTOS DA
UNIVERSIDADE FEDERAL DA
BAHIA & HUPES/UFBA



Continuação do Parecer: 4.231.669

computador e ao arquivo banco de dados). Somente o pessoal autorizado (membros da equipe terão acesso aos dados e estes serão treinados em boas práticas médicas e de pesquisa e assinarão os termos de anuência de sigilo). Os formulários de pesquisa serão guardados em locais seguros.

Descarte do material: os questionários e os arquivos de dados serão destruídos conforme a legislação vigente no momento ou 5 anos após a publicação dos dados de encerramento do protocolo.

Benefícios:

Os resultados obtidos com este estudo poderão ajudar os médicos a entender melhor a infecção por Coronavírus, identificar pacientes que podem evoluir com maior gravidade e com isso melhorar o tratamento desses pacientes. Pode não obter benefício direto nesse estudo, mas, esse conhecimento pode ajudar futuros pacientes. Se houver algum resultado que possa beneficiar os pacientes ele será comunicado imediatamente. Publicidade dos resultados: Os dados da pesquisa serão tornados públicos, sejam eles favoráveis ou não.

Comentários e Considerações sobre a Pesquisa:

Vide conclusões.

Considerações sobre os Termos de apresentação obrigatória:

A folha de rosto anexada não possui os dados e assinaturas do agente financiador.

O TCLE é apresentado e atende às determinações estabelecidas pela Resolução 466/12 do CNS. Deixa claro por sua vez que a dispensa de TCLE informada diz respeito somente a etapa do estudo retrospectivo e que os pacientes arrolados para a coorte prospectiva terão Termos de Consentimento Livre e Esclarecido.

Recomendações:

No item "Grupos em que serão divididos os participantes da pesquisa neste centro" presente no documento de informações básicas, o pesquisador informa datas para realização da etapa prospectiva divergentes do cronograma apresentado e descrita em outras partes do protocolo. Recomenda-se a revisão.

A folha de rosto apresentada não consta dados e assinatura do agente financiador. A despeito de

Endereço: Rua Augusto Viana, s/nº - 1º Andar
Bairro: Canela **CEP:** 40.110-060
UF: BA **Município:** SALVADOR
Telefone: (71)3283-8043 **Fax:** (71)3283-8140 **E-mail:** cep.hupes@ebserh.gov.br

UFBA - HOSPITAL
UNIVERSITÁRIO PROF.
EDGARD SANTOS DA
UNIVERSIDADE FEDERAL DA
BAHIA & HUPES/UFBA



Continuação do Parecer: 4.231.669

ter carta de compromisso. Recomenda-se a revisão e ajuste.

Conclusões ou Pendências e Lista de Inadequações:

O projeto apresenta relevância científica e potencial de contribuir para o melhor acompanhamento e diagnóstico de pacientes com Covid-10. Não foram encontrados óbices éticos.

Diante do exposto, emitimos parecer de APROVAÇÃO.

Considerações Finais a critério do CEP:

Aprovado conforme parecer da relatora.

O participante da pesquisa tem a liberdade de recusar-se a participar ou de retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo ao seu cuidado (Res. CNS 466/12) e deve receber uma via do Termo de Consentimento Livre e Esclarecido, na íntegra, completamente assinado.

O pesquisador deve desenvolver a pesquisa conforme delineada no protocolo aprovado e descontinuar o estudo somente após análise das razões da descontinuidade pelo CEP que o aprovou, aguardando seu parecer, exceto quando perceber risco ou dano não previsto ao sujeito participante ou quando constatar a superioridade de regime oferecido a um dos grupos da pesquisa que requeiram ação imediata.

O CEP deve ser informado de todos os efeitos adversos ou fatos relevantes que alterem o curso normal do estudo. É papel do pesquisador assegurar medidas imediatas adequadas frente a evento adverso grave ocorrido (mesmo que tenha sido em outro centro) e enviar notificação ao CEP e à Agência Nacional de Vigilância Sanitária – ANVISA – junto com seu posicionamento.

Eventuais modificações ou emendas ao protocolo devem ser apresentadas ao CEP de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas.

Relatórios parciais e final devem ser apresentados ao CEP, inicialmente em ____/____/____ e ao término do estudo.

Endereço: Rua Augusto Viana, s/nº - 1º Andar
 Bairro: Canela CEP: 40.110-060
 UF: BA Município: SALVADOR
 Telefone: (71)3283-8043 Fax: (71)3283-8140 E-mail: cep.hupes@ebserh.gov.br

**UFBA - HOSPITAL
UNIVERSITÁRIO PROF.
EDGARD SANTOS DA
UNIVERSIDADE FEDERAL DA
BAHIA & HUPES/UFBA**



Continuação do Parecer: 4.231.669

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1578849.pdf	14/08/2020 18:20:28		Aceito
Declaração de Pesquisadores	Declaracao_HUAC_EBSERH_UFCG.pdf	14/08/2020 16:05:03	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Declaracao_Anuencia_HUPES.pdf	14/08/2020 16:04:05	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_UFJF.pdf	14/08/2020 16:02:18	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Carta_Anuencia_UFCG_01Jul2020.pdf	14/08/2020 16:01:40	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_STELA.pdf	14/08/2020 16:01:25	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_SAME_HUPES.pdf	14/08/2020 16:01:03	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_RADTEC.pdf	14/08/2020 16:00:51	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_HUAC_EBSERH_UFCG.pdf	14/08/2020 16:00:34	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_do_servico_HUPES.pdf	14/08/2020 16:00:09	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_Bioimagem.pdf	14/08/2020 15:59:50	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_Compromisso_Thamires_Rosado_Reina.pdf	14/08/2020 15:56:27	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_Compromisso_Rafaela_Argonnes.jpeg	14/08/2020 15:56:14	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_Compromisso_Margarida_Celia_Lima_Costa_Neves.pdf	14/08/2020 15:55:55	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_compromisso_Marcos_Machado.pdf	14/08/2020 15:55:05	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_compromisso_Jose_Todesco.pdf	14/08/2020 15:54:49	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_Compromisso_Eduardo_Martins_Netto.pdf	14/08/2020 15:54:28	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_compromisso_Denilson_Sell.pdf	14/08/2020 15:54:04	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_Compromisso_Danillo_Menezes_Oliveira.pdf	14/08/2020 15:53:38	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_compromisso_Cesar_Araujo.jpg	14/08/2020 15:53:00	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_Compromisso_Carlos_Brites.pdf	14/08/2020 15:52:44	Nadja Sales Pacheco	Aceito

Endereço: Rua Augusto Viana, s/nº - 1º Andar

Bairro: Canela

CEP: 40.110-060

UF: BA

Município: SALVADOR

Telefone: (71)3283-8043

Fax: (71)3283-8140

E-mail: cep.hupes@ebserh.gov.br

**UFBA - HOSPITAL
UNIVERSITÁRIO PROF.
EDGARD SANTOS DA
UNIVERSIDADE FEDERAL DA
BAHIA & HUPES/UFBA**



Continuação do Parecer: 4.231.669

Declaração de Pesquisadores	Termo_de_compromisso_Bruno_Santana.pdf	14/08/2020 15:50:28	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_Compromisso_Arquimedes_Aires_Braga_de_Lira.pdf	14/08/2020 15:50:12	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_compromisso_Aran_Bey_Tcholakian_Morales.pdf	14/08/2020 15:49:46	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_Compromisso_Antonio_Carlos_Lemos.pdf	14/08/2020 15:49:25	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Termo_de_Compromisso_Andreia_de_Sousa_Lessa.pdf	14/08/2020 15:49:03	Nadja Sales Pacheco	Aceito
Outros	Questionario_versao_1_26_de_junho_2020.pdf	14/08/2020 15:46:49	Nadja Sales Pacheco	Aceito
Outros	Questionario_versao_1_26_de_junho_2020.docx	14/08/2020 15:46:33	Nadja Sales Pacheco	Aceito
Outros	Carta_de_Encaminhamento_de_Projeto_ao_CEP.pdf	14/08/2020 15:45:11	Nadja Sales Pacheco	Aceito
Outros	Carta_de_Encaminhamento_de_Projeto_ao_CEP.docx	14/08/2020 15:44:29	Nadja Sales Pacheco	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_versao1_de_26_de_junho_de_2020.pdf	14/08/2020 15:42:10	Nadja Sales Pacheco	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_versao1_de_26_de_junho_de_2020.docx	14/08/2020 15:41:59	Nadja Sales Pacheco	Aceito
Projeto Detalhado / Brochura Investigador	PROJETO_TCeIA_versao_1_de_26_de_junho_de_2020.pdf	14/08/2020 15:41:45	Nadja Sales Pacheco	Aceito
Projeto Detalhado / Brochura Investigador	PROJETO_TCeIA_versao_1_de_26_de_junho_de_2020.docx	14/08/2020 15:41:30	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Declaracao_de_Curriculos_da_equipe_do_estudo.pdf	14/08/2020 15:40:44	Nadja Sales Pacheco	Aceito
Orçamento	Orcamento_projeto_de_pesquisa_versao_1_26JUN2020.pdf	14/08/2020 15:40:08	Nadja Sales Pacheco	Aceito
Orçamento	Orcamento_projeto_de_pesquisa_versao_1_26JUN2020.docx	14/08/2020 15:39:54	Nadja Sales Pacheco	Aceito
Declaração de Pesquisadores	Declaracao_de_Curriculos_da_equipe_do_estudo.docx	14/08/2020 15:38:09	Nadja Sales Pacheco	Aceito
Declaração de Instituição e Infraestrutura	Declaracao_de_Infraestrutura_RADTEC.pdf	14/08/2020 15:37:18	Nadja Sales Pacheco	Aceito
Cronograma	Cromograma_projeto_de_pesquisa_versao_1_26JUN2020.pdf	14/08/2020 15:37:02	Nadja Sales Pacheco	Aceito

Endereço: Rua Augusto Viana, s/nº - 1º Andar

Bairro: Canela

CEP: 40.110-060

UF: BA

Município: SALVADOR

Telefone: (71)3283-8043

Fax: (71)3283-8140

E-mail: cep.hupes@ebserh.gov.br

UFBA - HOSPITAL
UNIVERSITÁRIO PROF.
EDGARD SANTOS DA
UNIVERSIDADE FEDERAL DA
BAHIA e HUPES/UFBA



Continuação do Parecer: 4.231.669

Cronograma	Cromograma_projeto_de_pesquisa_vers ao_1_26JUN2020.docx	14/08/2020 15:36:47	Nadja Sales Pacheco	Aceito
Folha de Rosto	Folha_de_Rosto_13AGO2020.pdf	13/08/2020 15:09:24	Nadja Sales Pacheco	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

SALVADOR, 24 de Agosto de 2020

Assinado por:

NATANAEL MOURA TEIXEIRA DE JESUS
(Coordenador(a))

Endereço: Rua Augusto Viana, s/nº - 1º Andar

Bairro: Canela

CEP: 40.110-060

UF: BA

Município: SALVADOR

Telefone: (71)3283-8043

Fax: (71)3283-8140

E-mail: cep.hupes@ebserh.gov.br

UFCG - HOSPITAL
UNIVERSITÁRIO ALCIDES
CARNEIRO DA UNIVERSIDADE
FEDERAL DE CAMPINA
GRANDE / HUAC - UFCG



PARECER CONSUBSTANCIADO DO CEP

Elaborado pela Instituição Coparticipante

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Desenvolvimento e validação de software para diagnóstico de covid-19 baseado em Tomografia Computadorizada, Radiômica e Inteligência Artificial.

Pesquisador: Eduardo Martins Netto

Área Temática:

Versão: 1

CAAE: 36624620.0.3003.5182

Instituição Proponente: Hospital Universitário Alcides Carneiro - Campina Grande/PB

Patrocinador Principal: FINANCIADORA DE ESTUDOS E PROJETOS - FINEP

DADOS DO PARECER

Número do Parecer: 4.818.271

Apresentação do Projeto:

De acordo com o pesquisador o pesquisador solicita nessa emenda as seguintes alterações: 1. Mudança de investigadores membro da equipe; 2. Trocar instituição que foi com endereço errado (Alcides Carneiro); 3. Incluir Instituição; 4. Acrescentar texto na metodologia do projeto; 5. Alterar o número da amostra de 160 para 180 participantes. Trata-se de um protocolo para o desenvolvimento de um software RIS/PACS com aplicação de radiômica e inteligência artificial que tenha capacidade de auxiliar no diagnóstico de Covid-19 em relação a outras doenças em pacientes com sintomas gripais. Projeto multicêntrico, do tipo coorte retrospectiva, com 10 pacientes de cada instituição com diagnósticos relacionados à pneumonia anterior a outubro-2019, e mais 10 pacientes de cada instituição com diagnósticos relacionados à Covid-19 entre abril-2020 e junho-2020. Total de pacientes 80 de participantes retrospectivo. Apresenta como desfecho primário uma inovação incremental ao processo de diagnóstico por imagens TC, induzida por uma forte inovação de produtos RIS/PACS ao incorporar técnicas de radiômica e inteligência artificial. O advento da radiômica e da inteligência artificial representa uma quebra de paradigma na medicina, pois tem possibilitado melhorias importantes nos modelos diagnósticos de várias doenças, assim como vem contribuindo com a abordagem de medicina personalizada. Embora alguns padrões da TC sejam mais comuns em pacientes com Covid-19 em relação à

Endereço: Rua: Dr. Carlos Chagas, s/ n

Bairro: São José

CEP: 58.107-670

UF: PB

Município: CAMPINA GRANDE

Telefone: (83)2101-5545

Fax: (83)2101-5523

E-mail: cep@huac.ufcg.edu.br

UFCG - HOSPITAL
UNIVERSITÁRIO ALCIDES
CARNEIRO DA UNIVERSIDADE
FEDERAL DE CAMPINA
GRANDE / HUAC - UFCG



Continuação do Parecer: 4.818.271

outras infecções, padrões inespecíficos representam uma limitação da TC. O emergente mercado da radiômica e inteligência artificial como ferramentas para apoio ao diagnóstico pode ser uma forte aliada no combate à Covid-19, pois desvendam informações relevantes sobre a doença, imperceptíveis ao olho humano. Este projeto visa o desenvolvimento de um software para quantificação de imagens e inteligência artificial para o diagnóstico do Covid-19, embarcado em um sistema RIS/PACS na nuvem.

Objetivo da Pesquisa:

O pesquisador destaca como objetivos:

Objetivo Primário:

Otimizar o diagnóstico do Covid-19 por tomografia computadorizada a partir de melhoramento tecnológico de um software RIS/PACS com aplicação de radiômica e inteligência artificial.

Objetivo Secundário:

- Desenvolver um aplicativo para segmentação automática de regiões de interesse (lesões de Covid-19) de imagens de tomografia computadorizada, embarcado em um sistema PACS (Picture Archiving and Communication System);
- Desenvolver um software para realizar a quantificação radiômica (padrões de imagens) de lesões pulmonares do Covid-19;
- Desenvolver painel de mineração de dados clínicos (achados dos laudos e fatores de risco) para aplicação em modelos preditivos;
- Desenvolver escores de risco usando dados clínicos, padrões de radiômica da tomografia computadorizada e inteligência artificial (IA)

Avaliação dos Riscos e Benefícios:

O pesquisador descreve como riscos e benefícios:

Riscos: Os riscos associados com o armazenamento das imagens pode ser a perda de confidencialidade, porém elas estarão sob sigilo e com código sob acesso somente dos pesquisadores. O desenvolvimento do trabalho implica em riscos mínimos de perda de confidencialidade para os participantes da pesquisa. Estes riscos serão minimizados utilizando-se códigos de acesso nos computadores. Os dados clínicos serão tratados como sigilosos. Os questionários serão codificados. Uma lista mestra ligará os nomes e identificação dos pacientes aos questionários de pesquisa (que usarão códigos). Portanto, a identidade dos pacientes será

Endereço: Rua: Dr. Carlos Chagas, s/ n
Bairro: São José **CEP:** 58.107-670
UF: PB **Município:** CAMPINA GRANDE
Telefone: (83)2101-5545 **Fax:** (83)2101-5523 **E-mail:** cep@huac.ufcg.edu.br

**UFCG - HOSPITAL
UNIVERSITÁRIO ALCIDES
CARNEIRO DA UNIVERSIDADE
FEDERAL DE CAMPINA
GRANDE / HUAC - UFCG**



Continuação do Parecer: 4.818.271

preservada. Todo o cuidado será tomado quanto ao sigilo dos dados (senhas de acesso: ao computador e ao arquivo banco de dados). Somente o pessoal autorizado (membros da equipe terão acesso aos dados e estes serão treinados em boas práticas médicas e de pesquisa e assinarão os termos de anuência de sigilo). Os formulários de pesquisa serão guardados em locais seguros. Descarte do material: os questionários e os arquivos de dados serão destruídos conforme a legislação vigente no momento ou 5 anos após a publicação dos dados de encerramento do protocolo.

Benefícios: Os resultados obtidos com este estudo poderão ajudar os médicos a entender melhor a infecção por Coronavírus, identificar pacientes que podem evoluir com maior gravidade e com isso melhorar o tratamento desses pacientes. Pode não obter benefício direto nesse estudo, mas, esse conhecimento pode ajudar futuros pacientes. Se houver algum resultado que possa beneficiar os pacientes ele será comunicado imediatamente. **Publicidade dos resultados:** Os dados da pesquisa serão tornados públicos, sejam eles favoráveis ou não

Comentários e Considerações sobre a Pesquisa:

A pesquisa denota relevância científica por propor uma inovação incremental ao processo de diagnóstico por imagens TC, induzida por uma forte inovação de produtos RIS/PACS ao incorporar técnicas de radiômica e IA.

Considerações sobre os Termos de apresentação obrigatória:

Foram anexados ao sistema todos os documentos de apresentação obrigatória.

Conclusões ou Pendências e Lista de Inadequações:

Somos de parecer favorável a aprovação dessa emenda.

Considerações Finais a critério do CEP:

Liberado Ad Referendum

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Projeto Detalhado / Brochura Investigador	Protocolo_v4_adendo.docx	07/05/2021 15:53:10	José Alisson Araújo Nunes	Aceito

Endereço: Rua: Dr. Carlos Chagas, s/ n
Bairro: São José **CEP:** 58.107-670
UF: PB **Município:** CAMPINA GRANDE
Telefone: (83)2101-5545 **Fax:** (83)2101-5523 **E-mail:** cep@huac.ufcg.edu.br

**UFCG - HOSPITAL
UNIVERSITÁRIO ALCIDES
CARNEIRO DA UNIVERSIDADE
FEDERAL DE CAMPINA
GRANDE / HUAC - UFCG**



Continuação do Parecer: 4.818.271

TCLE / Termos de Assentimento / Justificativa de Ausência	DispensaTCLE.pdf	14/04/2021 21:19:03	Marcos Antônio Dórea Machado	Aceito
Outros	Carta_Encaminhamento_Abr2021.pdf	14/04/2021 16:48:20	Marcos Antônio Dórea Machado	Aceito
Outros	Protocolo_de_pesquisa_v4_adendo.pdf	14/04/2021 16:47:58	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___GEP.pdf	14/04/2021 16:36:46	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___CadastroPesquisadorThamirisRReina.pdf	14/04/2021 16:36:27	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___AnuenciaSuperintendencia.pdf	14/04/2021 16:36:10	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___Cadastro.pdf	14/04/2021 16:35:33	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___CurriculoThamiris.pdf	14/04/2021 16:35:21	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___DeclaracaoInfraestrutura.pdf	14/04/2021 16:34:45	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___DispensaTCLE.pdf	14/04/2021 16:34:15	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___planilhaOrçamentoFinanceiro.pdf	14/04/2021 16:34:01	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___ProjetoDePesquisa.pdf	14/04/2021 16:33:23	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___TermoDeConfidencialidadeSigilo.pdf	14/04/2021 16:32:47	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___DeclaracaoComprometimento.pdf	14/04/2021 16:32:26	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___CartaDeEncaminhamento.pdf	14/04/2021 16:27:10	Marcos Antônio Dórea Machado	Aceito
Outros	Carta.pdf	23/03/2021 12:32:16	Marcos Antônio Dórea Machado	Aceito
Outros	CrtCEPHUPESemenda1_9_12_2020.pdf	09/12/2020 11:14:17	Eduardo Martins Netto	Aceito
Outros	CrtCEPHUPESemenda1_9_12_2020.docx	09/12/2020 11:13:58	Eduardo Martins Netto	Aceito
Outros	Carta_Anuencia_UFJF.pdf	14/08/2020 16:02:18	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_STELA.pdf	14/08/2020 16:01:25	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_SAME_HUPES.pdf	14/08/2020 16:01:03	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_RADTEC.pdf	14/08/2020	Nadja Sales	Aceito

Endereço: Rua: Dr. Carlos Chagas, s/ n

Bairro: São José

CEP: 58.107-670

UF: PB

Município: CAMPINA GRANDE

Telefone: (83)2101-5545

Fax: (83)2101-5523

E-mail: cep@huac.ufcg.edu.br

**UFCG - HOSPITAL
UNIVERSITÁRIO ALCIDES
CARNEIRO DA UNIVERSIDADE
FEDERAL DE CAMPINA
GRANDE / HUAC - UFCG**



Continuação do Parecer: 4.818.271

Outros	Carta_Anuencia_RADTEC.pdf	16:00:51	Pacheco	Aceito
Outros	Carta_Anuencia_HUAC_EBSERH_UFCG.pdf	14/08/2020 16:00:34	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_do_servico_HUPES.pdf	14/08/2020 16:00:09	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_Bioimagem.pdf	14/08/2020 15:59:50	Nadja Sales Pacheco	Aceito
Outros	Questionario_versao_1_26_de_junho_2020.pdf	14/08/2020 15:46:49	Nadja Sales Pacheco	Aceito
Outros	Questionario_versao_1_26_de_junho_2020.docx	14/08/2020 15:46:33	Nadja Sales Pacheco	Aceito
Outros	Carta_de_Encaminhamento_de_Projeto_ao_CEP.pdf	14/08/2020 15:45:11	Nadja Sales Pacheco	Aceito
Outros	Carta_de_Encaminhamento_de_Projeto_ao_CEP.docx	14/08/2020 15:44:29	Nadja Sales Pacheco	Aceito
Projeto Detalhado / Brochura Investigador	PROJETO_TCeIA_versao_1_de_26_de_junho_de_2020.pdf	14/08/2020 15:41:45	Nadja Sales Pacheco	Aceito
Projeto Detalhado / Brochura Investigador	PROJETO_TCeIA_versao_1_de_26_de_junho_de_2020.docx	14/08/2020 15:41:30	Nadja Sales Pacheco	Aceito

Situação do Parecer:

Aprovado

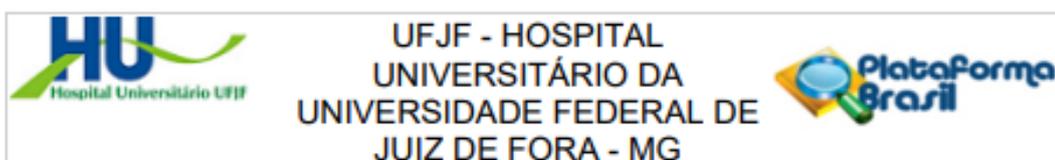
Necessita Apreciação da CONEP:

Não

CAMPINA GRANDE, 30 de Junho de 2021

Assinado por:
Andréia Oliveira Barros Sousa
(Coordenador(a))

Endereço: Rua: Dr. Carlos Chagas, s/ n
Bairro: São José **CEP:** 58.107-670
UF: PB **Município:** CAMPINA GRANDE
Telefone: (83)2101-5545 **Fax:** (83)2101-5523 **E-mail:** cep@huac.ufcg.edu.br



PARECER CONSUBSTANCIADO DO CEP

Elaborado pela Instituição Coparticipante

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Desenvolvimento e validação de software para diagnóstico de covid-19 baseado em Tomografia Computadorizada, Radiômica e Inteligência Artificial.

Pesquisador: Eduardo Martins Netto

Área Temática:

Versão: 1

CAAE: 36624620.0.3005.5133

Instituição Proponente: UNIVERSIDADE FEDERAL DE JUIZ DE FORA UFJF

Patrocinador Principal: FINANCIADORA DE ESTUDOS E PROJETOS - FINEP

DADOS DO PARECER

Número do Parecer: 4.926.688

Apresentação do Projeto:

Apresentação do projeto está clara, detalhada de forma objetiva, descreve as bases científicas que justificam o estudo, de acordo com as atribuições definidas na Resolução CNS 466/12 de 2012, item III. Descreve o projeto:

*As áreas de radiômica (do inglês, radiomics) e inteligência artificial (IA) têm recebido grande atenção nos últimos anos, devido à sua capacidade em melhorar modelos diagnóstico e definir fenótipos de indivíduos que podem melhor responder a determinado tratamento, trazendo uma grande contribuição na era da medicina personalizada. Em um contexto de pandemia da Covid-19, estes métodos têm sido objeto de debate.

Alguns estudos têm demonstrado potencial para a aplicação da radiômica em tomografias computadorizadas (TC) utilizando IA como método para diagnóstico do Covid-19. Plataformas de radiômica são restritas a alguns softwares livres capazes de quantificar padrões de imagens das lesões, mas com muitas limitações de uso prático pois precisam estar inseridos num workflow de interação com os médicos, e posteriormente servir de dados de entrada em outra aplicação com IA. Algumas empresas têm surgido na Europa e Estados Unidos disponibilizando soluções de radiômica. Estas empresas têm recebido investimentos em um nicho com oportunidades crescentes em várias áreas da medicina. Nosso projeto visa criar uma alternativa a estas soluções,

Endereço: Rua Catulo Breviglieri, s/n

Bairro: Santa Catarina

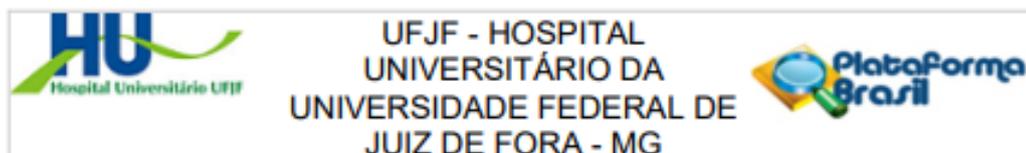
CEP: 36.036-110

UF: MG

Município: JUIZ DE FORA

Telefone: (32)4009-5217

E-mail: cep.hu@ufjf.edu.br



Continuação do Parecer: 4.926.688

uma vez que inexistem soluções nacionais com foco nesta área, a partir de uma plataforma desenvolvida inicialmente para o combate à Covid-19, mas escalável para outras áreas da medicina. Neste projeto, recursos humanos e tecnológicos da empresa Radtec, de quatro hospitais e outros parceiros tecnológicos serão compartilhados para apoiar a criação, a implantação e a consolidação de uma nova tecnologia para apoio inteligente ao diagnóstico de Covid19 em um projeto de inovação e empreendedorismo tecnológico.*

Objetivo da Pesquisa:

O Objetivo da pesquisa está bem delineado, apresenta clareza e compatibilidade com a proposta, tendo adequação da metodologia aos objetivos pretendido, de acordo com as atribuições definidas na Norma Operacional CNS 001 de 2013, item 3.4.1 - 4.

O projeto é descrito da seguinte forma:

*Objetivos: Otimizar o diagnóstico do Covid-19 por tomografia computadorizada a partir de melhoramento tecnológico de um software RIS/PACS com aplicação de radiômica e inteligência artificial:

- Desenvolver um aplicativo para segmentação automática de regiões de interesse (lesões de covid-19) de imagens de tomografia computadorizada, embarcado em um Sistema de Informação Radiológica (Radiology Information System - RIS) / Sistema de Comunicação e Arquivamento de Imagens (Picture Archiving and Communication System - PACS);
- Desenvolver um software para realizar a quantificação radiômica (padrões de imagens) de lesões pulmonares do Covid-19;
- Desenvolver painel de mineração de dados clínicos (achados dos laudos e fatores de risco) para abastecer outras aplicações de análises.
- Desenvolver escores de risco para assistência inteligente do diagnóstico da Covid-19, usando dados clínicos, padrões de radiômica da tomografia computadorizada e inteligência artificial (IA).*

Avaliação dos Riscos e Benefícios:

Identificação dos riscos e as possibilidades de desconfortos e benefícios esperados estão adequadamente descritos. A avaliação dos Riscos e Benefícios estão de acordo com as atribuições definidas na Resolução CNS 466/12 de 2012, itens III; III.2 e V.

Descreve o projeto:

*Os riscos associados estão relacionados com o armazenamento das imagens e dos dados clínicos, podendo ocorrer a perda de confidencialidade, porém elas estarão sob sigilo e com código sob acesso somente dos pesquisadores.

Endereço: Rua Catulo Breviglieri, s/n	CEP: 36.036-110
Bairro: Santa Catarina	
UF: MG	Município: JUIZ DE FORA
Telefone: (32)4009-5217	E-mail: cep.hu@ufjf.edu.br



Continuação do Parecer: 4.926.688

Os resultados obtidos com este estudo poderão ajudar os médicos a entender melhor a infecção por Coronavírus, identificar pacientes que podem evoluir com maior gravidade e com isso melhorar o tratamento desses pacientes. Os pacientes da pesquisa podem não obter benefício direto nesse estudo, mas esse conhecimento pode ajudar futuros pacientes.*

Comentários e Considerações sobre a Pesquisa:

O projeto está bem estruturado, delineado e fundamentado, sustenta os objetivos do estudo em sua metodologia de forma clara e objetiva, e se apresenta em consonância com os princípios éticos norteadores da ética na pesquisa científica envolvendo seres humanos elencados na resolução 466/12 do CNS e com a Norma Operacional Nº 001/2013 CNS.

O presente projeto já fora analisado em CEP de origem e pela CONEP, sendo a presente proposta uma emenda para inclusão do HU-UFJF como local de pesquisa. Os dados dos projetos estão de acordo com a proposta já analisada anteriormente.

Considerações sobre os Termos de apresentação obrigatória:

O protocolo de pesquisa está em configuração adequada, apresenta FOLHA DE ROSTO devidamente preenchida, com o título em português, identifica o patrocinador pela pesquisa, estando de acordo com as atribuições definidas na Norma Operacional CNS 001 de 2013 item 3.3 letra a; e 3.4.1 item 16. Apresenta o TERMO DE DISPENSA DO TCLE de acordo com a Resolução CNS 466 de 2012, item: IV.8. O Pesquisador apresenta titulação e experiência compatível com o projeto de pesquisa, estando de acordo com as atribuições definidas no Manual Operacional para CPEs. Apresenta DECLARAÇÃO de infraestrutura e de concordância com a realização da pesquisa de acordo com as atribuições definidas na Norma Operacional CNS 001 de 2013 item 3.3 letra h. Por se tratar de emenda, os documentos analisados visam verificar a anuência de infraestrutura local, bem como se os documentos juntados com os dados do centro local reproduzem os dados e metodologias já analisadas e aprovadas pelo centro de origem e CONEP.

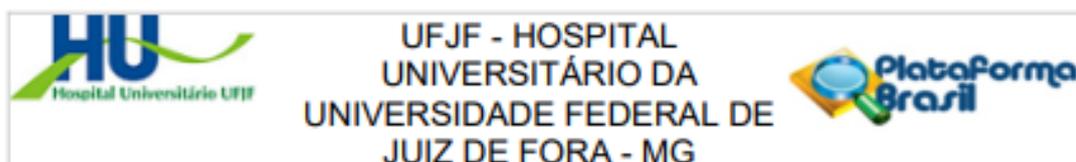
Contudo: o cronograma apresentado necessita ser adequado por não estar de acordo com a Norma Operacional CNS Nº 001 de 2013, item 3.4.1.9., visto que o mesmo prevê a inclusão e análise dos dados é anterior a presente análise do CDEP.

Conclusões ou Pendências e Lista de Inadequações:

Pendente

De acordo com a Norma Operacional CNS Nº 001 de 2013, item 3.4.1.9., "Todos os protocolos de pesquisa devem conter, obrigatoriamente: (...) Cronograma: informando a duração total e as diferentes etapas da pesquisa, em número de meses, com compromisso explícito do pesquisador

Endereço: Rua Catulo Breviglieri, s/n	
Bairro: Santa Catarina	CEP: 36.036-110
UF: MG	Município: JUIZ DE FORA
Telefone: (32)4009-5217	E-mail: cep.hu@ufjf.edu.br



Continuação do Parecer: 4.926.888

de que a pesquisa somente será iniciada a partir da aprovação pelo Sistema CEP-Conep*. Sendo assim, o cronograma não pode indicar o início do estudo em data que antecede a tramitação da análise ética.

Considerações Finais a critério do CEP:

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
TCLE / Termos de Assentimento / Justificativa de Ausência	DispensaTCLE.pdf	14/04/2021 21:19:03	Marcos Antônio Dórea Machado	Aceito
Outros	Carta_Encaminhamento_Abr2021.pdf	14/04/2021 16:48:20	Marcos Antônio Dórea Machado	Aceito
Outros	Protocolo_de_pesquisa_v4_adendo.pdf	14/04/2021 16:47:58	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___GEP.pdf	14/04/2021 16:36:46	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___CadastroPesquisadorThamirisRReina.pdf	14/04/2021 16:36:27	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___AnuenciaSuperintendencia.pdf	14/04/2021 16:36:10	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___Cadastro.pdf	14/04/2021 16:35:33	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___CurriculoThamiris.pdf	14/04/2021 16:35:21	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___DeclaracaoInfraestrutura.pdf	14/04/2021 16:34:45	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___DispensaTCLE.pdf	14/04/2021 16:34:15	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___planilhaOrcamentoFinanceiro.pdf	14/04/2021 16:34:01	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___ProjetoDePesquisa.pdf	14/04/2021 16:33:23	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___TermodeConfidencialidadeSigilo.pdf	14/04/2021 16:32:47	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___DeclaracaoComprometimento.pdf	14/04/2021 16:32:26	Marcos Antônio Dórea Machado	Aceito
Outros	HUUFJF___CartadeEncaminhamento.pdf	14/04/2021 16:27:10	Marcos Antônio Dórea Machado	Aceito
Outros	Carta.pdf	23/03/2021 12:32:16	Marcos Antônio Dórea Machado	Aceito

Endereço: Rua Catulo Breviglieri, s/n

Bairro: Santa Catarina

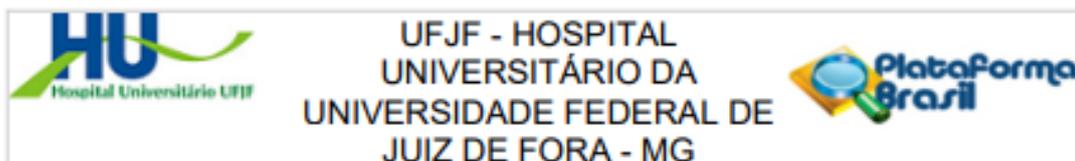
CEP: 36.036-110

UF: MG

Município: JUIZ DE FORA

Telefone: (32)4009-5217

E-mail: cep.hu@uff.edu.br



Continuação do Parecer: 4.926.688

Outros	CriCEPHUPESemenda1_9_12_2020.pdf	09/12/2020 11:14:17	Eduardo Martins Netto	Aceito
Outros	CriCEPHUPESemenda1_9_12_2020.docx	09/12/2020 11:13:58	Eduardo Martins Netto	Aceito
Outros	Carta_Anuencia_UFJF.pdf	14/08/2020 16:02:18	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_STELA.pdf	14/08/2020 16:01:25	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_SAME_HUPES.pdf	14/08/2020 16:01:03	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_RADTEC.pdf	14/08/2020 16:00:51	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_HUAC_EBSERH_UFCG.pdf	14/08/2020 16:00:34	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_do_servico_HUPES.pdf	14/08/2020 16:00:09	Nadja Sales Pacheco	Aceito
Outros	Carta_Anuencia_Bioimagem.pdf	14/08/2020 15:59:50	Nadja Sales Pacheco	Aceito
Outros	Questionario_versao_1_26_de_junho_2020.pdf	14/08/2020 15:46:49	Nadja Sales Pacheco	Aceito
Outros	Questionario_versao_1_26_de_junho_2020.docx	14/08/2020 15:46:33	Nadja Sales Pacheco	Aceito
Outros	Carta_de_Encaminhamento_de_Projeto_ao_CEP.pdf	14/08/2020 15:45:11	Nadja Sales Pacheco	Aceito
Outros	Carta_de_Encaminhamento_de_Projeto_ao_CEP.docx	14/08/2020 15:44:29	Nadja Sales Pacheco	Aceito
Projeto Detalhado / Brochura Investigador	PROJETO_TCeIA_versao_1_de_26_de_junho_de_2020.pdf	14/08/2020 15:41:45	Nadja Sales Pacheco	Aceito
Projeto Detalhado / Brochura Investigador	PROJETO_TCeIA_versao_1_de_26_de_junho_de_2020.docx	14/08/2020 15:41:30	Nadja Sales Pacheco	Aceito

Situação do Parecer:

Pendente

Necessita Apreciação da CONEP:

Não

Endereço: Rua Catulo Breviglieri, s/n

Bairro: Santa Catarina

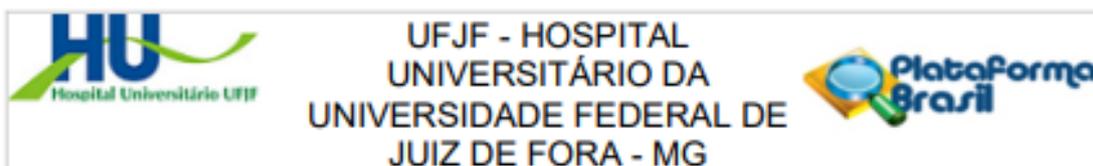
CEP: 36.036-110

UF: MG

Município: JUIZ DE FORA

Telefone: (32)4009-5217

E-mail: cep.hu@uff.edu.br



Continuação do Parecer: 4.926.688

JUIZ DE FORA, 24 de Agosto de 2021

Assinado por:
João Beccon de Almeida Neto
(Coordenador(a))

Endereço: Rua Catulo Breviglieri, s/n
Bairro: Santa Catarina **CEP:** 36.036-110
UF: MG **Município:** JUIZ DE FORA
Telefone: (32)4009-5217 **E-mail:** oep.hu@ufjf.edu.br