



**UNIVERSIDADE FEDERAL DA BAHIA
FACULDADE DE MEDICINA DA BAHIA**

**PROGRAMA DE PÓS-GRADUAÇÃO
EM MEDICINA E SAÚDE**



MARCOS ANTÔNIO DÓREA MACHADO

**TECNOLOGIAS DA INFORMAÇÃO E MÉTODOS
COMPUTACIONAIS PARA GERENCIAMENTO, OTIMIZAÇÃO E
MEDICINA DE PRECISÃO EM DEPARTAMENTOS DE IMAGENS
MÉDICAS**

TESE DE DOUTORADO

SALVADOR

2023



**UNIVERSIDADE FEDERAL DA BAHIA
FACULDADE DE MEDICINA DA BAHIA**

**PROGRAMA DE PÓS-GRADUAÇÃO
EM MEDICINA E SAÚDE**



MARCOS ANTÔNIO DÓREA MACHADO

**TECNOLOGIAS DA INFORMAÇÃO E MÉTODOS
COMPUTACIONAIS PARA GERENCIAMENTO, OTIMIZAÇÃO E
MEDICINA DE PRECISÃO EM DEPARTAMENTOS DE IMAGENS
MÉDICAS**

Tese apresentada ao Programa de Pós-graduação em Medicina e Saúde, da Faculdade de Medicina da Bahia, Universidade Federal da Bahia, como requisito para a obtenção do grau de Doutor em Medicina e Saúde.

Orientador: Eduardo Martins Netto

Salvador

2023

Ficha Catalográfica elaborada pelo Sistema Universitário de Bibliotecas (SIBI/UFBA),
com dados fornecidos pelo(a) autor(a)

Machado, Marcos Antônio Dórea

Tecnologias da Informação e Métodos Computacionais para Gerenciamento, Otimização e Medicina de Precisão em Departamentos de Imagens Médicas/ Marcos Antônio Dórea Machado. —

-- Salvador, 2023.

113 f. : il.

Orientador: Prof. Dr. Eduardo Martins Netto.

Tese (Doutorado - Programa de Pós-graduação em Medicina e Saúde) -- Universidade Federal da Bahia, Faculdade de Medicina da Bahia, 2023

Inclui referências e anexos.

1. Tecnologia médica. 2. padronização. 3. protocolos. 4. biomarcadores;
5. inteligência artificial.

MARCOS ANTÔNIO DÓREA MACHADO

**TECNOLOGIAS DA INFORMAÇÃO E MÉTODOS
COMPUTACIONAIS PARA GERENCIAMENTO, OTIMIZAÇÃO E
MEDICINA DE PRECISÃO EM DEPARTAMENTOS DE IMAGENS
MÉDICAS**

Tese apresentada ao Programa de Pós-graduação em Medicina e Saúde, da Faculdade de Medicina da Bahia, Universidade Federal da Bahia, como requisito para a obtenção do grau de Doutor em Medicina e Saúde.

Orientador: Eduardo Martins Netto

Salvador, 13 de janeiro de 2023.

BANCA EXAMINADORA

André Rodrigues Durães

Doutor em Medicina e Saúde
Universidade Federal da Bahia

Lucas de Oliveira Vieira

Doutor em Cardiologia
Escola Bahiana de Medicina e Saúde Pública

Felipe Alves Mourato

Doutor em Biologia Aplicada à Saúde
Universidade Federal de Pernambuco

Susana Souza Lalic

Doutora em Física
Universidade Federal de Sergipe

Ana Maria Marques da Silva

Doutora em Física
Pontifícia Universidade Católica do Rio Grande do Sul

AGRADECIMENTOS

Agradeço a Deus, pela saúde e oportunidade de ter colocado em meu caminho pessoas incríveis que contribuíram para a realização deste projeto. À minha família, especialmente minha esposa Tiale Acrux, pelo amor, compreensão e apoio. Aos colegas da Radtec, Bruno Santana, Cleiton Queiroz e Vinícius Menezes, pelas contribuições em todas as etapas deste trabalho, pelas discussões científicas, contribuição nos experimentos e participação nos projetos e empreendimentos científicos e tecnológicos. A Dayane Ponce, pelo apoio no desenvolvimento de novos produtos tecnológicos e pela sua capacidade de transferir o conhecimento gerado em aplicações práticas. Agradeço também à Dra Simone Brandão, Dra Nadja Rolim, Dr André Gustavo, Dr Augusto Ximenes, Dr Lucas Vieira pelos ensinamentos da prática médica, especialmente da medicina nuclear e sistemas de informações para a medicina. Ao Monte-Tabor, por acreditar e viabilizar nossos esforços em oferecer o melhor diagnóstico e tratamento para os pacientes em todas as suas necessidades. Ao meu orientador Prof. Dr. Eduardo Netto, pelos ensinamentos sobre medicina baseada em evidências, epidemiologia e métodos diagnósticos, pelas discussões e contribuições nas produções científicas e tecnológicas. À Prof. Dra Adriana Latado, Prof. Dra Liliane Lins-Kusterer e demais professores do PPGMS que contribuíram para o meu aprendizado. Ao Hospital Universitário Prof. Edgard Santos/Ebserh por permitir a realização deste trabalho. Não poderia deixar de agradecer também aos apoiadores deste trabalho: o Conselho Nacional de Pesquisas (CNPq), Fundação de Amparo à Pesquisa do Estado da Bahia (Fapesb), à Financiadora de estudos e projetos (Finep) e a Agência Internacional de Energia Atômica (IAEA). Por fim, quero dedicar esta obra ao meu amado filho Gabriel Acrux Machado, que o bom Deus lhe permita usufruir dos benefícios da ciência.

“Deixar o mundo como ele é, é nos humilharmos; torná-lo melhor é agir como filhos de Deus. Somos pessoas que trabalham dia a dia, com dedicação, porque compreendemos que o resgate da humanidade é, também, de nossa responsabilidade. Devemos dar tudo o que somos e tudo que temos”. Don Luigi Maria Verzé

RESUMO

A área da saúde demanda por tecnologias da informação e métodos computacionais para melhorar a produtividade dos serviços e oferecer assistência personalizada aos pacientes. Este trabalho buscou desenvolver e explorar sistemas e métodos computacionais para implementar melhorias no gerenciamento, otimizar os exames, e acessar novos biomarcadores e assinaturas com inteligência artificial para apoio à decisão. Foram desenvolvidos e implementados *softwares* com o conceito *Workflow Based Approach* (WBA) e métodos computacionais escritos em linguagem Python para melhorar a gestão e otimizar os protocolos de exames. Um workflow para acesso a novos biomarcadores e assinaturas IA foi desenvolvido e validado em pacientes com CT de COVID-19, 18F-FDG-PET/CT de câncer de colo do útero e 18F-FDG-PET/CT de linfoma Hodgkin. O workflow demonstrou-se válido em análises de robustez: repetitividade (erro < 5%), reprodutibilidade (coeficiente de correlação intraclasse, ICC > 90%) e correlação clínica ($p < 0,05$). Os modelos preditivos para 18F-FDG-PET/CT de câncer de colo do útero e 18F-FDG-PET/CT de linfoma Hodgkin apresentaram desempenho geral de AUC=0,74 e AUC=0,96, respectivamente. Um novo software que utiliza métodos de IA para apoio ao diagnóstico da COVID-19 em CT de tórax de pacientes com pneumonia foi disponibilizado e validado em um PACS/Viewer. Sem apoio do software, os médicos tiveram desempenho médio de 83,4% de sensibilidade, e 64,3% de especificidade. Com o apoio do software, o desempenho melhorou para 87,1% de sensibilidade, e 91,1% de especificidade. Adicionalmente, o software melhorou a concordância entre observadores, de moderado para substancial, em uma escala construída a partir do coeficiente de concordância Cohen's Kappa.

Palavras-chave: Tecnologia médica, padronização, protocolos, biomarcadores, inteligência artificial, radiologia

ABSTRACT

The healthcare industry demands information technologies and computational methods to improve productivity and offer personalized assistance to patients. This work aimed to develop and explore information systems and computational methods to improve the management, optimize the exams, and provide new biomarkers and artificial intelligence signatures for decision making in medical imaging facilities. We developed softwares based on *Workflow Based Approach* (WBA) concept and computational methods using Python language to improve the management and optimize the exam protocols. A framework to access new biomarkers and artificial intelligence (AI) signatures was developed and validated in COVID-19 CT patients, 18F-FDG-PET/CT cervical cancer and 18F-FDG-PET/CT Hodgkin lymphoma patients. This framework was feasible for robustness analysis: repetitivity (error < 5%), reproducibility (intraclass correlation coefficient, ICC > 90%) and clinical correlation ($p < 0.05$). The overall performance of predictive models was AUC=0.74 and AUC=0.96 for 18F-FDG-PET/CT cervical cancer and 18F-FDG-PET/CT Hodgkin lymphoma, respectively. A new AI software to support thorax CT COVID-19 diagnoses was implemented and validated within PACS/Viewer. Without the support of software, physicians performed with mean sensitivity and specificity of 83.4% and 64.3%, respectively. When they were assisted with AI software, mean sensitivity and specificity were 87.1% and 91.1%, respectively. In addition, AI software improved the inter-rater reliability from moderate to substantial agreement in a Cohen's Kappa scale.

Keywords: medical technology, standardization, protocols, biomarkers, artificial intelligence, radiology

LISTA DE ILUSTRAÇÕES

Figura 1 - Box plot dos valores de ruído (CV) para as reconstruções OSEM e PSF do fabricante Siemens.....	17
Figura 2 – Esquema de posologias do 18F-FDG por tipo de reconstrução, normalizado para 3 min de tempo de imagem em cada região do corpo (<i>bed position</i>).	18
Figura 3 - Etapas para extração de biomarcadores de imagens	25
Figura 4 – Usabilidade e arquitetura para diagnóstico inteligente da COVID-19.....	26
Figura 5 – Quadro resumo das etapas do estudo.....	27

LISTA DE ABREVIATURAS E SIGLAS

AI	<i>Artificial Intelligence</i>
API	<i>Application Programing Interface</i> (Interface de Programação de Aplicativos)
CT	<i>Computed tomography</i>
DICOM	<i>Digital Imaging and Communications in Medicine</i>
FDG	Fluorodeoxyglucose
HU	<i>Hounsfield Unit</i> (Unidade Hounsfield, ou número CT)
HL7	<i>Health Level 7</i>
IA	Inteligência Artificial
INPI	Instituto Nacional de Propriedade Industrial
KDD	<i>Knowledge Discovery</i> (Descoberta de Conhecimento)
LGPD	Lei Geral de Proteção de Dados
MR	<i>Magnetic Resonance</i>
MTV	<i>Metabolic Tumor Volume</i> (Volume Metabólico do Tumor)
NLP	<i>Natural Language Processing</i> (Processamento de Linguagem Natural)
OSEM	<i>Ordered Subset Expectation Maximization</i>
PACS	<i>Picture Archiving and Communication System</i> (Sistema de Comunicação e Arquivamento de Imagens)
P&D	Pesquisa e desenvolvimento
PET	<i>Positron emission tomography</i> (Tomografia por emissão de pósitron)
PET/CT	<i>Positron emission tomography (hybrid)</i> (Equipamento híbrido de tomografia por emissão de positron)
PSF	<i>Point Spread Function</i>
RM	Ressonância Magnética
ROC	<i>Receiver Operating Characteristic</i> (Característica operacional do receptor)
ROI	<i>Region of interest</i> (Região de interesse)
SPECT	<i>Single Photon Emission Computed Tomography</i> (Tomografia por emissão de fóton único)
SUV	<i>Standard uptake value</i> (Valor de captação padrão)
TC	Tomografia computadorizada
TLG	<i>Total Lesion Glycolisis</i> (Glicólise Total da Lesão)
TRL	<i>Technology Level Readiness</i> (Nível de Maturidade Tecnológica)
VOI	<i>Volume of Interest</i> (Volume de Interesse)

SUMÁRIO

1. INTRODUÇÃO	11
2. OBJETIVOS	13
3. REVISÃO DE LITERATURA	14
3.1 Sistemas de Informações.....	14
3.1.1 Sistemas RIS.....	14
3.1.2 Sistemas de Arquivamento e Comunicação de Imagens (PACS)	14
3.1.3 Interoperabilidade.....	15
3.2 Padronização de Protocolos de Exames	15
3.3 Otimização de protocolos de imagens	16
3.4 Biomarcadores de imagens.....	19
3.4.1 Linfoma de Hodgkin	20
3.4.2 Câncer de colo de útero	20
3.4.3 COVID-19.....	21
4. MATERIAL E MÉTODOS.....	22
4.1 Software e método para gerenciamento de filas de exames	23
4.2 Software e método para otimização do tempo de imagens	24
4.3 Padronização, seleção e análise de biomarcadores	24
4.4 Workflow para acesso a novos biomarcadores de imagens	24
4.5 <i>Software</i> de IA para diagnóstico da COVID-19	26
5. RESULTADOS	11
5.1 ARTIGOS ORIGINAIS.....	11
6. DISCUSSÃO	70
7. CONCLUSÃO	73
8. PERSPECTIVAS.....	74
REFERÊNCIAS	93
ANEXOS	100

1. INTRODUÇÃO

O avanço dos computadores e das metodologias para desenvolvimento de sistemas têm possibilitado a incorporação de novas tecnologias na medicina de uma forma rápida e crescente. Esta transformação ocorre em diferentes áreas e linhas de cuidado, desde novas soluções de gestão hospitalar, soluções de telemedicina e tecnologias com inteligência artificial para apoio à decisão médica.

Várias especialidades na área de saúde, especialmente os serviços de radiologia e medicina nuclear enfrentam dificuldades relacionadas à complexa variabilidade dos processos, riscos quanto às medicações e aumento crescente de procedimentos mais complexos. A falta de ferramentas para gestão das filas de atendimento leva a ineficiências operacionais (ex.: aumento do tempo de espera do paciente) em todas as etapas da jornada do cuidado, desde o acesso ao sistema de saúde, passando pelo atendimento (ex.: recepção, salas de exames, priorização de laudos), até sua liberação médica. Devido à carência de soluções de gestão sob medida, a maioria dos serviços ainda utilizam sistemas genéricos que não contemplam as especificidades da área (ex: procedimentos realizados em múltiplas etapas e múltiplos protocolos clínicos), permitindo novos caminhos para a aplicação de tecnologias da informação. O conceito de *Workflow Based Approach* (WBA) se destaca na adoção de *softwares* de gestão sob medida, que se baseia na modelagem dos processos de saúde, tal que o sistema consiga gerenciar as rotinas contemplando as especificidades de cada processo. Além de otimizar a gestão, o sistema com WBA permite também agregar novas funcionalidades de *software* específicas para determinado fluxo de trabalho.

Se, por um lado, sistemas de gestão mais sofisticados podem suprimir a baixa produtividade dos médicos - que gastam bastante tempo para coletar dados de pacientes, tomar decisões e elaborar os relatórios médicos - por outro lado, a introdução de novos métodos computacionais pode auxiliar os médicos acessar novas informações para a tomada de decisões. É o caso do emprego de técnicas para extração de biomarcadores de imagens (denominado radiômica), que permite acessar informações de textura tumoral, densidade, atividade metabólica, dentre outras características das imagens médicas. O surgimento da pandemia da COVID-19 afetou não só as pessoas como também a capacidade de atendimento do sistema de saúde. A COVID-19 causa uma pneumonia cuja diferenciação da etiologia no exame de tomografia computadorizada não é possível, por enquanto. Com isso, métodos computacionais

com radiômica e inteligência artificial podem facilitar a diferenciação diagnóstica e, portanto, auxiliar na melhoria da gestão da saúde.

A Inteligência Artificial (IA) tem ganhado notoriedade na área da saúde nos últimos anos, em parte pelo alto potencial para melhoramento da classificação de doenças, contribuindo com a medicina de precisão. As aplicações têm sido diversas, desde combinações de variáveis (ex.: pressão arterial, diabetes) para criação de modelos com *machine learning*; até aplicações complexas que envolvem processamento de imagens com *deep learning* para aceleração de imagens, como Ressonância Magnética (RM), Tomografia por Emissão de Pósitrons (PET) e Tomografia por Emissão de Fóton Único (SPECT) que podem otimizar o tempo da realização do exame, reduzindo de cerca de 20min para 5-10min.

Existem atualmente aplicações para aceleração de exames e para extração de biomarcadores que são, em geral, parte dos equipamentos de aquisição de imagens. Porém, para utilizar estas soluções os usuários devem adquirir licenças junto ao fabricante dos equipamentos. Isto limita, sobretudo, a incorporação dos biomarcadores em modelos preditivos com IA, devido aos desafios de interoperabilidade. A implementação de soluções de IA em um sistema WBA independente dos equipamentos de aquisição de imagens poderia facilitar o acesso a estas novas tecnologias. Por outro lado, o desenvolvimento de *softwares* para aceleração de imagens requer a incorporação de técnicas que simulam imagens de diferentes padrões de qualidade, a fim de verificar a consistência dos *softwares* em transformar imagens de baixo padrão de qualidade em imagens clínicas convencionais. Neste mesmo sentido, o estudo de novos biomarcadores (extração, reprodutibilidade, correlação clínica) precede à disponibilização dos modelos de apoio à decisão com IA, o que demanda um continuum de desenvolvimento tecnológico e análises até a validação clínica.

Este trabalho de pesquisa e desenvolvimento (P&D) traz soluções em termos de melhorias na gestão de serviços de diagnóstico por imagens utilizando recursos de *software* (menor tempo de exame, menor radiação e segurança do paciente), acesso a novos biomarcadores, e algoritmos de inteligência artificial para apoio ao diagnóstico.

2. OBJETIVOS

Geral:

Explorar a aplicação de sistemas e métodos computacionais para melhorias na gestão e na assistência aos pacientes em serviços de diagnóstico por imagens.

Específicos:

- Desenvolver e implementar metodologias e *software* para gerenciamento e otimização dos exames de imagens: tempo de imagem, menor dose de radiação e segurança do paciente.
- Desenvolver e implementar metodologias para extração, seleção e análise de biomarcadores de imagens.
- Desenvolver e implementar algoritmos preditivos em exames de PET e CT para o diagnóstico de linfoma de Hodgkin, câncer de colo de útero e COVID-19.

3. REVISÃO DE LITERATURA

Neste capítulo são abordadas as fundamentações teóricas das principais tecnologias digitais, sistemas de informações e métodos computacionais aplicados na medicina, com ênfase nos *softwares* usados no segmento de diagnóstico por imagens.

3.1 Sistemas de Informações

3.1.1 Sistemas RIS

Os departamentos de radiologia estão entre os primeiros serviços médicos a adotar sistemas eletrônicos de registro na década de 1960-70, operando de forma relativamente independente do hospital, quando passaram a ser chamados de Sistemas de Informações em Radiologia (do inglês, *Radiology Information System* – RIS). As principais rotinas que foram inicialmente digitalizadas se referem à identificação do paciente (e da solicitação de exames) na base de dados do hospital e a monitoração do paciente durante a imagem e o laudo médico (ASH and BATES *et al.*, 2005). Desde então, os *softwares* RIS passaram por evoluções ao se integrar em sistemas PACS (do inglês, *Picture Archiving and Communication System* – PACS) na década de 1990 e incorporar fluxos de trabalho mais eficientes, permitindo realizar mais estudos, laudos elaborados mais rápidos e disponibilizados quase que imediatamente (NANCY *et al.*, 2013).

A variabilidade dos processos em saúde dificulta a implantação de sistemas de informações, que devem ser generalizáveis para diferentes tipos de procedimentos médicos. Neste sentido, os *softwares* de gestão deveriam incorporar conhecimento, a fim de melhor auxiliar os usuários na gestão dos processos, no que se denomina WBA (do inglês, *Workflow Based Approach*), permitindo maior flexibilidade e número de funcionalidades na gestão de processos em saúde (ZANG *et al.*, 2009).

3.1.2 Sistemas de Arquivamento e Comunicação de Imagens (PACS)

Os sistemas PACS são aplicações que armazenam e distribuem imagens médicas no padrão DICOM (do inglês, *Digital Imaging and Communications in Medicine*). Estes sistemas são integrados com os equipamentos de aquisição de imagens como tomógrafos, ultrassons, dentre outras modalidades, e quando integrados com sistemas de visualização de imagens, permitem a análise dos exames de forma remota. Desde o surgimento destes sistemas na década

de 1980, o segmento de diagnóstico por imagens passou por uma rápida evolução, sendo adotado o conceito, em 1995, de **PACS de Larga Escala** as aplicações que se enquadrassem, dentre outros critérios, o gerenciamento de mais de 20.000 exames por ano (MANSOORI *et al.*, 2012). Os altos custos envolvidos na implantação de um sistema PACS suscitaram dúvidas sobre o custo-efetividade destes sistemas. Em 2009, AYAL e colegas publicaram um estudo demonstrando que o benefício econômico na integração dos sistemas hospitalares com os RIS e PACS alcançam 63% ao final dos primeiros 12 meses pós-implantação (AYAL and SEIDMAN *et al.*, 2009). Os ganhos com eficiência e produtividade incluem o acesso aos exames através da internet, eliminação de processos manuais redundantes (ex: *worklist* nas estações de trabalho que integram com a agenda do RIS) e aumento da produtividade médica.

3.1.3 Interoperabilidade

A interoperabilidade se refere à capacidade de diferentes *softwares* trabalharem em conjunto, garantindo que os sistemas interajam para trocar informações de maneira eficaz. Como vimos, as tecnologias das informações são fundamentais para a melhoria da assistência aos pacientes e para a otimização operacional dos serviços de saúde. Entretanto, apesar dos benefícios trazidos pela interoperabilidade entre sistemas, esta tarefa é muito difícil ou inalcançável em alguns casos. Isto se dá principalmente pela diversidade e heterogeneidade das diferentes ferramentas digitais, métodos, processos e procedimentos que resultam em muitos softwares proprietários sem terminologias padronizadas (IROJU *et al.*, 2013). A integração baseada em padrões de interoperabilidade é um dos aspectos de *softwares* para saúde mais modernos. Os principais padrões de interoperabilidade são os protocolos de comunicação DICOM (para imagens) e HL7 (para dados dos prontuários) (BLAZONA and KONCAR, 2007).

3.2 Padronização de Protocolos de Exames

Construir e manter um conjunto de protocolos consistentes e padronizados é dificultoso principalmente quando se tem múltiplos departamentos integrados em um grande sistema de saúde. Nestes casos, a falta de uniformidade nos protocolos assistenciais pode aumentar em 2,5 vezes a quantidade de protocolos de exames, como foi demonstrado no estudo de SACHS *et al.* (2017) com protocolos de CT e RM em 5 hospitais operando em rede. Neste estudo, os autores demonstraram uma experiência na padronização dos pedidos médicos e na realização dos

exames, reduzindo a quantidade de protocolos em aproximadamente 60%, usando uma metodologia baseada na indicação clínica do exame.

Metodologias de acreditação e políticas de qualidade são outro importante recurso para a padronização dos procedimentos, sobretudo quando estão envolvidos biomarcadores de imagens usados como escores para reportar os achados (KAALEP *et al.*, 2018). Neste contexto, a padronização de biomarcadores é um passo importante para avançar os sistemas aplicados à radiologia e imagem molecular. Para isto, os protocolos de imagens devem ser estabelecidos no sentido da padronização do preparo, etapas, reconstrução de imagens, metodologias para comunicação dos resultados e garantia da qualidade dos processos (ACR, 2022; PADHANI, A. and OLLIVIER *et al.*, 2001; WAHL *et al.*, 2009; HOEKSTRA *et al.*, 2014).

3.3 Otimização de protocolos de imagens

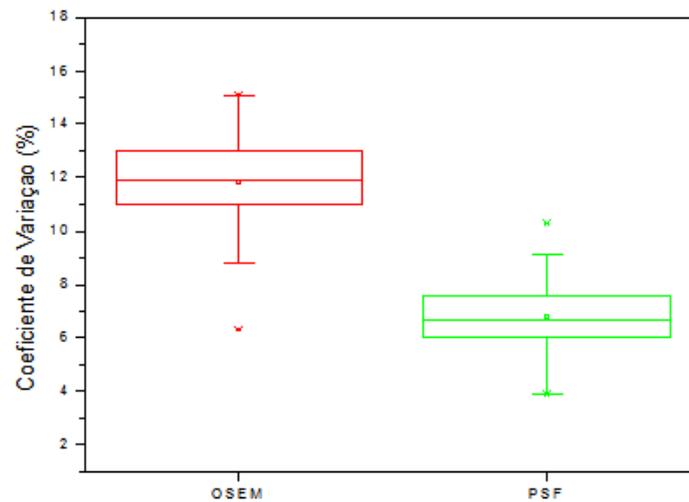
As imagens de PET, SPECT e RM adotam protocolos de aquisição de dados que podem priorizar o menor tempo de exames considerando um liminar de padrão de qualidade (KOOPEMAN *et al.*, 2016; MENEZES *et al.*, 2016; LECCHI, M. *et al.*, 2017; JAMBOR, 2017). Nestas modalidades, o tempo de aquisição das imagens é um fator crucial na produtividade do serviço de saúde, onde o tempo pode ser otimizado com a incorporação de detectores mais sensíveis nos equipamentos e/ou com *softwares* de reconstrução tomográfica que permitem recuperar imagens realizadas com baixo sinal (maior relação sinal/ ruído) a fim de obter novas imagens com padrão diagnóstico (DESHMANE *et al.*, 2012; VALENTA *et al.*, 2010; MACHADO *et al.*, 2019).

Por outro lado, as imagens de CT são realizadas em alguns segundos, por isso a otimização dos protocolos desta modalidade tem foco na redução da dose de radiação aos pacientes. Por isso os equipamentos de CT mais modernos possuem tecnologia de modulação de corrente do tubo de raio-X, que regula a intensidade da radiação de acordo com a densidade da região do corpo, mantendo a qualidade das imagens e expondo o paciente à quantidade mínima de radiação necessária para um padrão de qualidade pré-determinado (MARTIN and SOOKPENG, 2016).

Em geral, as soluções de *software* são proprietárias do fabricante do equipamento de imagens, o que traz dificuldades em padronizar protocolos e metodologias entre diferentes modelos de equipamentos. A figura 1 ilustra como diferentes *softwares* de reconstrução de imagens geram padrões de qualidade (ruído) diferentes para o equipamento PET/CT. Esta implementação gera desempenhos que podem reduzir o tempo de imagens e/ou a quantidade de

radiofármacos utilizada no paciente (figura 2), mantendo-se padrões de detectabilidade e quantificação (MACHADO *et al* 2017).

Figura 1 - Box plot dos valores de ruído (CV) para as reconstruções OSEM e PSF do fabricante Siemens.

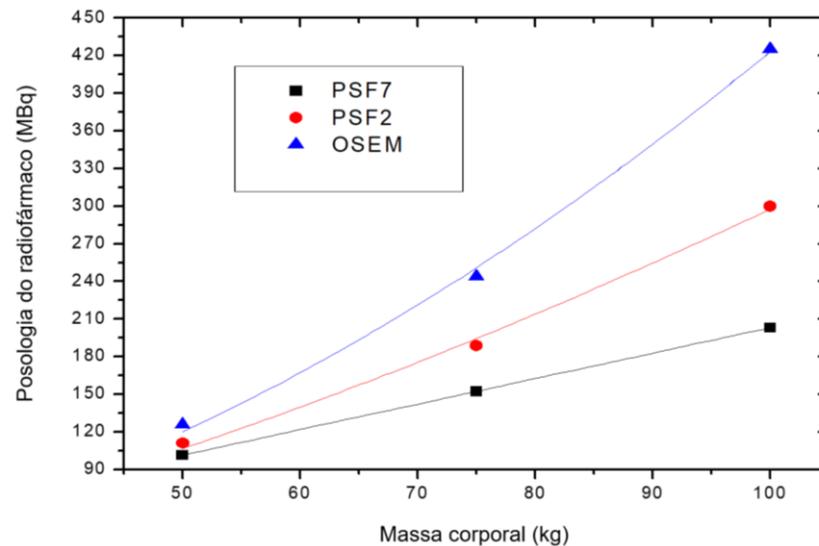


Adaptado de MACHADO *et al* 2017.

Comparação do ruído (CV) nas imagens entre os *softwares* de reconstruções *ordered subset expectation maximization* (OSEM) e *Point Spread Function* (PSF) (PET/CT Siemens), demonstrando a superioridade (menor ruído) da reconstrução PSF em relação à reconstrução OSEM. O menor ruído intrínseco possibilita realizar imagens mais rápidas e/ou com menor posologia do radiofármaco.

Na figura 2, é apresentada a comparação da posologia do radiofármaco necessária para cada tipo de *software* de reconstrução. Na PSF7 (reconstrução PSF com filtro de 7mm), gera quantificações equivalente à reconstrução OSEM com filtro de 5mm; na PSF2 (reconstrução PSF otimizada para melhoria da detectabilidade), gera quantificações em escalas diferentes em relação à PSF7 e OSEM.

Figura 2 – Esquema de posologias do ^{18}F -FDG por tipo de reconstrução, normalizado para 3 min de tempo de imagem em cada região do corpo (*bed position*).



Adaptado de MACHADO *et al* 2017.

Em 2017, nosso grupo incorporou no sistema Nuclearis o método para otimização da posologia do radiofármaco e o tempo da aquisição das imagens (MENEZES *et al*, 2016), de modo que os usuários do *software* pudessem configurar parâmetros dos equipamentos, padrões de qualidade, parâmetros de produção (posologia e tempo) e operar o sistema para otimizar os recursos de forma prática (MENEZES *et al*, 2018).

Com a popularização da inteligência artificial, outras soluções de *software* para aceleração da aquisição de imagens e/ou redução de doses de radiação, têm sido propostas usando redes neurais (método *deep-learning*) para supressão do ruído da imagem (*denoise*). Estes algoritmos são capazes de reconstruir o resultado de uma imagem diagnóstica a partir de um padrão de imagem de baixo sinal (alto ruído) (YANG *et al*, 2017; KAPLAN *et al*, 2019; SANAAT *et al*, 2021; ALMEIDA *et al*, 2021; OLIA *et al*, 2022). Estas soluções têm a vantagem de operar de forma independente do equipamento de aquisição de imagens. O equipamento pode gerar a imagem de baixa qualidade (com ruído) e enviar para o sistema PACS, a partir do qual outros softwares podem atuar. Outra vantagem destas soluções é que equipamentos de imagens e soluções mais antigas ainda em uso no mercado poderão se beneficiar destas soluções sem a necessidade de grandes investimentos em atualizações dos seus sistemas.

3.4 Biomarcadores de imagens

Os biomarcadores podem ser compreendidos como grandezas mensuráveis da condição de saúde, da severidade ou da presença de algum estado de doença (STRIMBU e TAVEL *et al*, 2010). As múltiplas modalidades de imagens como CT, RM, PET e SPECT são amplamente utilizadas para diagnóstico, estadiamento, prognóstico e monitoração da resposta ao tratamento. A capacidade de extrair escores como volume tumoral, fração de ejeção do ventrículo esquerdo ou mesmo a captação glicolítica do tumor reforça a importância destes métodos de imagens no manejo dos pacientes. Biomarcadores baseados na textura da imagem a partir do tecido, na morfologia e no metabolismo da substância injetada tem-se tornado possíveis de serem obtidos a partir de métodos computacionais para processamento de imagens, passando a fornecer informações complementares sobre o estado do paciente (COOK *et al*, 2014; VALLIÈRES *et al*, 2018). Este campo tem sido denominado de radiômica, onde vários autores reportaram achados importantes em várias doenças e modalidades de equipamentos (TIXIER *et al*, 2011; SUI *et al*, 2019; HUANG *et al*, 2021; PARK *et al*, 2020).

Um importante campo de pesquisas é a padronização e a reprodutibilidade destes biomarcadores (NAMIAS *et al*, 2018; CRANDALL *et al*, 2021; PFAEHLER *et al*, 2021). Estas iniciativas têm sido essenciais para disponibilizar ferramentas para qualificar e implementar novos biomarcadores na rotina clínica (ORLHAC *et al*, 2018, PARK *et al*, 2020; ZWANENBURG; *et al*, 2020).

Plataformas de radiômica e analytics são restritas a alguns softwares livres (ex: LIFEx, ImageJ) e outros poucos comerciais (ex: Radiomics, Sophia Genetics) capazes de quantificar padrões nas imagens das lesões. Em geral, estes sistemas são difíceis de ser incorporados no dia a dia da prática clínica porque não têm uma usabilidade prática, principalmente pela dificuldade de interoperabilidade com os demais sistemas hospitalares. Após quantificar os atributos radiômicos nestes sistemas, é necessário analisar quais podem se caracterizar como biomarcadores na prática, em termos de reprodutibilidade e associação clínica, servindo então de ferramenta de apoio à decisão.

Medicina de Precisão

A medicina de precisão é definida como um processo envolvendo a derivação de novas taxonomias baseadas em fenotipagem profunda. Para implementar a medicina de precisão, os resultados clínicos devem ser retroalimentados para obter ganho do conhecimento e assim,

aumentar a precisão (KOENIG et al, 2017). Neste contexto, os biomarcadores de imagens podem contribuir como informação incremental para o aprofundamento desta fenotipagem. De acordo com KOENIG e colaboradores (2017), são necessários esforços contínuos para alcançar ainda mais precisão e individualização, sendo que o ganho de conhecimento incremental, possivelmente de novas fontes de dados, pode incrementar ainda mais a acurácia.

3.4.1 Linfoma de Hodgkin

Nas últimas décadas, a terapia do linfoma Hodgkin evoluiu ao ponto de alcançar mais de 90% de sucesso, onde o exame 18F-FDG PET/CT se tornou o padrão ouro para o estadiamento e resposta ao tratamento, sobretudo como ferramenta para a personalização da quimioterapia e radioterapia (METTLER *et al*, 2018). O exame de 18F-FDG PET/CT tem a capacidade de detectar a extensão da doença e o comprometimento da medula óssea, permitindo o estadiamento de doença nodal e extra linfática de forma não invasiva. Adicionalmente à monitoração de mudanças durante o tratamento, alguns estudos apontam que o exame basal (antes do início do tratamento) tem o potencial de prever a resposta ao tratamento usando biomarcadores como o volume metabólico do tumor (COTTEREAU *et al*, 2018).

Os biomarcadores clássicos como a captação tumoral (SUV, do inglês *Standardized Uptake Value*) e volume tumoral (MTV) ignoram a heterogeneidade do tumor, que poderia servir como covariável para explicar o comportamento da doença. Milgron e colegas usaram um modelo de aprendizado de máquinas incluindo co-variáveis de textura e alcançaram aproximadamente 95% de acurácia, indicando que a radiômica tem potencial para incrementar a capacidade preditiva do comportamento da doença (MILGROM *et al*, 2019).

3.4.2 Câncer de colo de útero

O câncer de colo de útero é uma das causas de morte mais frequente entre as mulheres, especialmente em países em desenvolvimento (BRAY *et al*, 2018). Em pacientes com estágios mais avançados da doença, o exame 18F-FDG PET/CT é recomendado para avaliar o linfonodo e a doença metastática (FERLAY *et al*, 2015; SUNG *et al*, 2021). O biomarcador de captação tumoral (SUV) é o mais usado na técnica de PET/CT, informando sobre o metabolismo glicolítico do tumor. Quanto mais ávido pela glicose, mais agressivo é o tumor. Outros biomarcadores têm sido usados como o volume metabólico (MTV) e o produto entre a captação

e o volume (glicólise total, do inglês *Total Lesion Glycolisis* - TLG) (IM *et al*, 2018; KHIEWVAN *et al*, 2016). Dentre os vários biomarcadores disponíveis, aqueles relacionados à textura possuem alta correlação com a heterogeneidade do comportamento biológico (ex: necrose, hipóxia) e, portanto, poderiam servir como preditores de resposta.

3.4.3 COVID-19

A tomografia computadorizada (CT) é usada no diagnóstico e estadiamento da pneumonia desempenhando, portanto, um importante papel no manejo da pandemia da COVID-19 (KWEE *et al*, 2020). Entretanto, o padrão das lesões pulmonares vistos na CT não são patognomônicos, impondo uma limitação o método (BARBOSA *et al*, 2020). Com isso, alguns estudos mostraram o potencial de métodos com radiômica e IA para melhorar a especificidade da CT na pneumonia da COVID-19 em relação a outros tipos de pneumonia (SYEDA *et al*, 2021). Portanto, a incorporação de métodos de reconhecimento de padrões nos *softwares* de visualização de imagens pode trazer ganhos importantes no manejo da COVID-19.

4. MATERIAL E MÉTODOS

Com apoio do programa brasileiro de apoio a *startups* em 2014, nosso grupo iniciou o desenvolvimento de um *software* de gestão (denominado Nuclearis) com o conceito WBA, sobretudo para lidar com as especificidades das modalidades PET e SPECT, que usam medicamentos radioativos e fluxos de processos complexos (programa Startup Brasil CNPq, projeto 463765/2014, RADTEC, chamada CNPq/MCTI/SEPIN 12/2014 Programa Startup) (MACHADO *et al.*, 2016).

Em 2019, evoluímos a plataforma Nuclearis para incorporar uma solução PACS com laudos e prontuários eletrônicos estruturados integrados (FAPESB: Edital Bahia Inovação/FAPESB/SECTI N° 008/2016, projeto 19.571.128.5420, contrato n° 03.10.0322.00), hospedado na nuvem (*Amazon Web Services*), visando aumentar a produtividade da equipe multiprofissional a partir de ferramentas de gestão integradas.

Neste trabalho, foram desenvolvidos 1) *software* e método para gerenciamento de filas de exames, 2) *software* e método de particionamento de imagens para simulação de imagens SPECT com diferentes padrões de qualidade, facilitando estudos de otimização dos exames, além de serem precursores de modelos *deep-learning* para supressão de ruído - *denoise*; 3) métodos para extração, padronização, seleção e análises de novos biomarcadores, 4) solução integrada no sistema RIS/PACS-Nuclearis para diagnóstico da COVID-19, usando *deep-learning* para segmentação automática de imagens, e *machine-learning* para disponibilização de assinaturas IA com radiômica.

Devido à pandemia da COVID-9, que causou 230.452 mortes no Brasil em 2020 (FIOCRUZ, 2021), no ano seguinte foi iniciado o desenvolvimento tecnológico emergencial (6 a 12 meses) de soluções visando a melhoria no gerenciamento dos serviços de saúde (FINEP: Edital n° 03/2020, Soluções tecnológicas inovadoras para produtos, serviços e processos implementadas por Startups e Empresas de Base Tecnológica aplicadas ao ambiente de pandemia de COVID-19, Contrato: 03.20.0113.00) e soluções visando o diagnóstico da COVID-19 por imagens de CT e inteligência artificial (FINEP: Edital n° 02/2020, Soluções Inovadoras para o combate ao COVID-19, Contrato: 03.20.0156.00).

Para a disponibilização das aplicações tecnológicas na rotina clínica, foi utilizada a metodologia de desenvolvimento de produtos baseada em TRL (do inglês, *Technology Readiness Level*), identificando as etapas de pesquisa básica, desenvolvimento tecnológico, validação e industrialização (norma ISO 16290:2013). Alguns recursos disponíveis na

plataforma Nuclearis, como workflows clínicos, prontuários, PACS/Viewer e APIs, facilitam a coleta e distribuição de dados para validação rápida dos produtos (TRL7), incorporados nos seguintes componentes tecnológicos da plataforma:

- **Workflow Clínico** - Padronização de protocolos, monitoração dos processos, gerenciamento de filas e acesso facilitado a dados clínicos.
- **Estatística & Analytics** – Painel de apoio à tomada de decisões.

Este trabalho complementou estas soluções com novas ferramentas para gerenciamento de serviços de imagens, otimização de protocolos e acesso a novos biomarcadores e assinaturas com IA. A metodologia para desenvolvimento de *software* foi baseada nos princípios de Gerenciamento de Projetos propostos pelo BPM CBOOK (CBOOK, 2009). Os *softwares* desenvolvidos neste trabalho foram depositados no sistema de Peticionamento Eletrônico para registro de *software* no Instituto Nacional de Propriedade Industrial (INPI).

4.1 Software e método para gerenciamento de filas de exames

Para avaliar o impacto do *software* de gestão de filas, 17 (dezesete) os usuários de 06 (seis) centros diagnósticos foram convidados a responder um formulário comparando a experiência antes e depois da implantação do *software* na rotina de trabalho (TRL7). O formulário é composto por quatro perguntas, com opções de respostas sendo *importante* ou *indiferente*, para as duas primeiras perguntas (A e B); e opções de respostas sendo *melhorou*, *piorou* ou *indiferente* para as duas últimas perguntas (C e D).

Questionário:

- A) Como você avalia a importância de um software para auxiliar na rastreabilidade do paciente?
- B) Como você avalia a importância de um software para gerenciar as filas de atendimento?
- C) Na sua experiência com outros sistemas, em relação ao Nuclearis, como você avalia o impacto na segurança do paciente para a administração dos radiofármacos?
- D) Na sua opinião, quanto o Nuclearis auxilia no gerenciando de filas no atendimento em relação a outros sistemas?

4.2 Software e método para otimização do tempo de imagens

Nesta etapa viabilizamos o acesso a séries de imagens (do mesmo paciente) com diferentes padrões de qualidade, simulando situações de baixo sinal (maior ruído). Este aparato irá apoiar na definição de protocolos otimizados para aquisição de imagens médicas. Adicionalmente, poderá ser aplicado também para a criação de novas tecnologias para recuperação de ruído usando *deep-learning*, visando a aceleração de exames e/ou redução de doses de radiação (ALMEIDA *et al*, 2021; OLIA *et al*, 2022).

A metodologia desenvolvida para acesso a imagens de baixo sinal utilizou o equipamento de eletrocardiograma, usado em exames de imagens cardíacas, para capturar múltiplos do sinal gerado pelos equipamentos de imagens. Com estes múltiplos, pode-se simular exames realizados com diferentes espectros de sinal/ ruído (ARMSTRONG *et al*, 2012).

4.3 Padronização, seleção e análise de biomarcadores

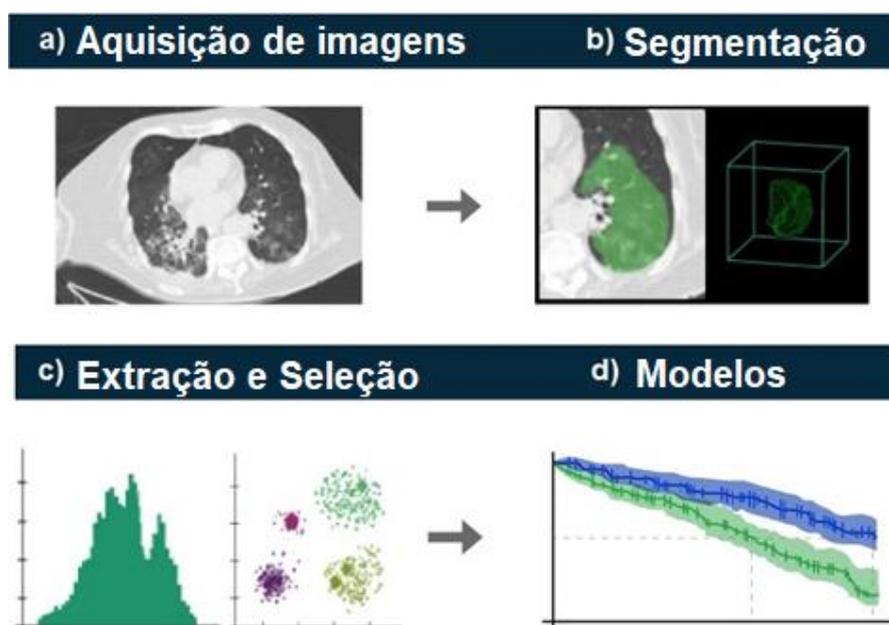
Esta etapa visou identificar, padronizar e avaliar a viabilidade de atributos radiômicos suficientemente robustos que poderão ser extraídos e usados na prática como biomarcadores. Obviamente, o acesso a estes atributos só fez sentido se eles complementavam uma informação existente (caracterizando-se como biomarcador), ou se são mais reprodutíveis ou de mais fácil acesso, considerando custo e disponibilidade. As metodologias para análise destes atributos radiômicos seguiram conforme proposto por ORLHAC (2018), ZWANENBURG (2020) e PFAEHLER (2021). Foram usados biomarcadores de PET para pacientes com linfoma, colo de útero e objetos simuladores de imagens para avaliações de reprodutibilidade e associação com desfechos clínicos (recidiva, morte), e foram usados biomarcadores CT para pacientes com pneumonias COVID e não-COVID para análises de padronização, seleção e harmonização de biomarcadores.

4.4 Workflow para acesso a novos biomarcadores de imagens

A figura 3 apresenta as etapas fundamentais que foram utilizadas para obtenção de biomarcadores radiômica. A etapa A, corresponde à imagem que o paciente gera no equipamento, com a indicação correta e com os protocolos de qualidade definidos. Tendo em vista que diferenças entre equipamentos (ex: modelo, softwares de reconstrução, filtros de

imagens) poderiam afetar a escala e a reprodutibilidade do biomarcador, tornou-se fundamental a determinação de parâmetros de qualidade nesta etapa do processo. A etapa B necessitou de algoritmos computacionais para delimitar a região ou órgão de interesse (segmentação). Quando a segmentação da imagem é feita de forma manual, introduz-se um fator operador-dependente no processo, pois muitos atributos radiômicos são sensíveis à região delimitada da imagem (ex: geometria da segmentação, região considerada, fatores de ponderação). Como consequência, é importante a escolha de biomarcadores que sejam invariáveis em função de pequenas variações na região segmentada. Preferencialmente, a segmentação automática deverá ser usada, pois elimina o viés do observador, além de automatizar o processo. Na etapa C foi feito o cálculo do atributo radiômico a partir das relações matemáticas entre os *pixels* (radiômica). Nesta etapa foram aplicados também métodos de seleção de atributos que sejam suficientemente robustos. Por exemplo, nesta etapa foram eliminados atributos com baixa reprodutibilidade e atributos que variavam pouco em função da condição clínica. Ao identificar atributos (biomarcadores) que eram estáveis e que representavam a doença, fez-se necessário eliminar aqueles que sejam fortemente correlacionados entre si, para simplificar o problema e evitar o viés de múltiplas hipóteses. A etapa D avaliou os biomarcadores juntos com outras condições clínicas do paciente para apoio à decisão médica em modelos multivariados.

Figura 3 - Etapas para extração de biomarcadores de imagens



O workflow desenvolvido nesta etapa permitiu maior rapidez na extração e seleção de biomarcadores. Nesta etapa, foram usadas as plataformas disponíveis ImageJ, LIFEx e

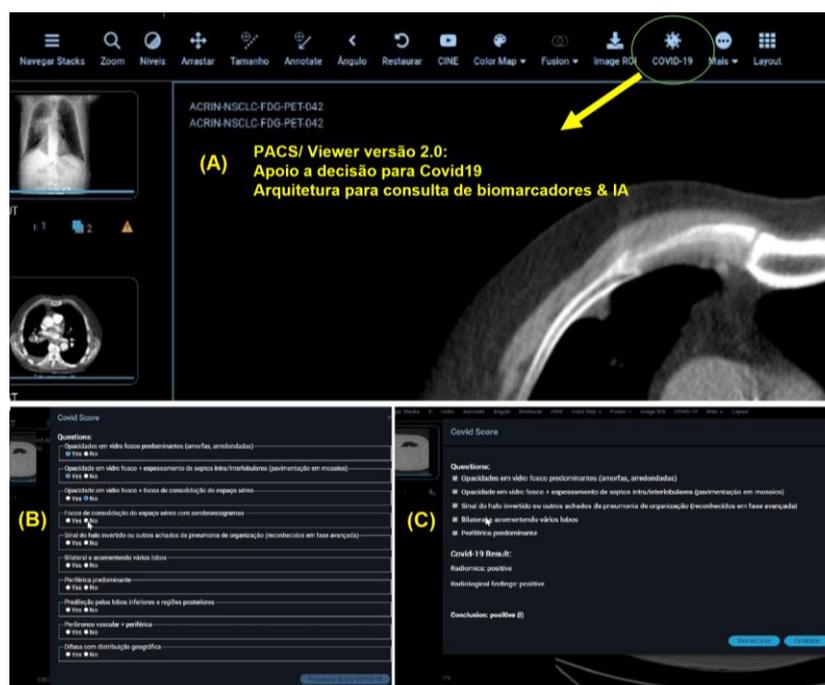
PyRadiomics como experiências para definir o conjunto de ferramentas mais adequado para o grupo (KANOUN *et al.*, 2015; NIOCHE *et al.*, 2018; GRIETHUYSEN *et al.*, 2017).

4.5 Software de IA para diagnóstico da COVID-19

Foram desenvolvidos componentes de *software* integrados no sistema RIS/PACS-Nuclearis, usando *deep-learning* para auto-segmentação de imagens, e *machine learning* para disponibilização de assinaturas IA com radiômica. Recursos com interoperabilidade DICOM foram desenvolvidos para acesso e operações com as imagens no conceito PACS Larga Escala (MANSOORI *et al.*, 2012) para incorporar a ferramenta de apoio à decisão para diagnóstico diferencial da COVID-19. A ferramenta foi construída usando dois escores com *machine learning*: a) questionários estruturados a partir de padronizações da taxonomia médica, e b) escores de biomarcadores de radiômica. Estes recursos computacionais foram desenvolvidos com a linguagem de programação Python (VANROSSUM *et al.*, 1995).

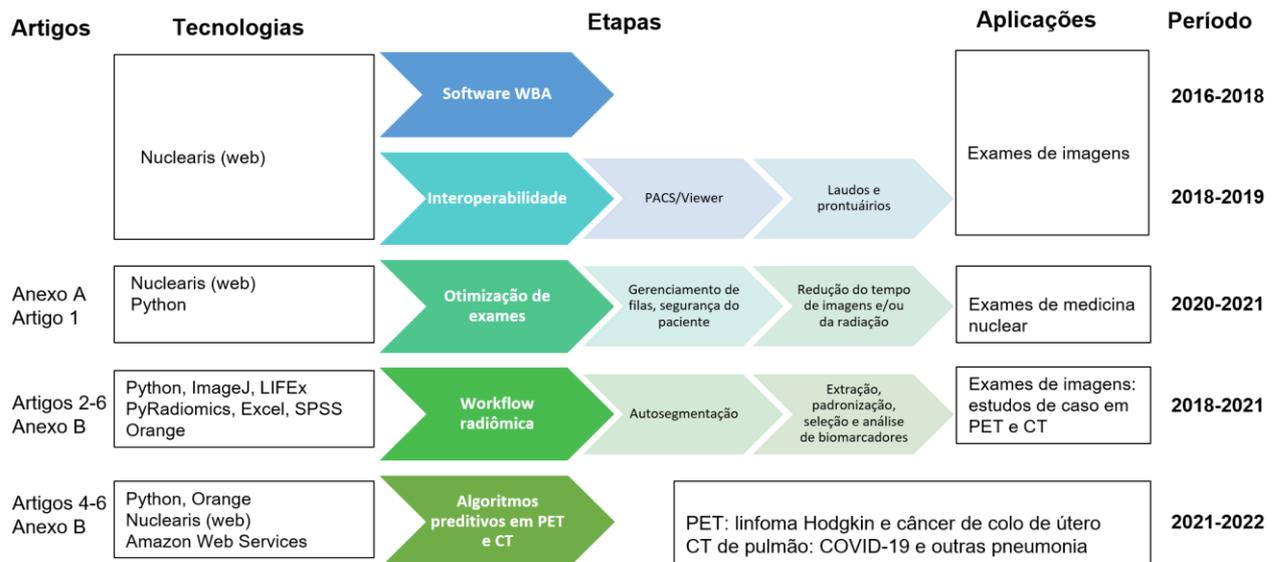
A figura 4 exemplifica o recurso que foi desenvolvido na plataforma para acionar o conjunto de inteligência para apoio ao diagnóstico da COVID-19. Em 4-A, temos um exame de CT de tórax de um paciente, demonstrando o botão para apoio inteligente ao diagnóstico da COVID-19. Em 4-B, O usuário escolhe 10 opções de um laudo estruturado, e na etapa final (4-C) o recurso retorna uma assinatura IA do laudo estruturado, e uma assinatura IA baseada em radiômica.

Figura 4 – Usabilidade e arquitetura para diagnóstico inteligente da COVID-19



A figura 5 apresenta um resumo das etapas, aplicações e ferramentas de *software*/ linguagens de programação usadas neste trabalho. O detalhamento metodológico de cada etapa pode ser verificado nos artigos apresentados a seguir.

Figura 5 – Quadro resumo das etapas do estudo



5. RESULTADOS

Os resultados correspondem à experiência dos usuários do software WBA, apresentado no Anexo A, aos artigos originais apresentados a seguir, e aos Anexos B e C.

5.1 ARTIGOS ORIGINAIS

Artigo nº 1

Technical note: Partitioning of gated single photon emission computed tomography raw data for protocols optimization

Journal of Applied Clinical Medical Physics

Publicado

TECHNICAL NOTE

Technical note: Partitioning of gated single photon emission computed tomography raw data for protocols optimization

Cleiton Cavalcante Queiroz^{1,2,3} | Marcos Antonio Dorea Machado^{1,2,4} |
Antonio Augusto Brito Ximenes² | Andre Gustavo Silva Pino³ |
Eduardo Martins Netto⁵

¹ Department of Nuclear Medicine, São Rafael Hospital/ Rede D'or, Salvador, Bahia, Brazil

² Department of Nuclear Medicine, Cardio Pulmonar Hospital/ Rede D'or, Salvador, Bahia, Brazil

³ Department of Nuclear Medicine, Hospital Universitario Professor Alberto Antunes/Ebserh, Maceio, Alagoas, Brazil

⁴ Department of Health Technology Evaluation, Complexo Hospitalar Universitário Prof. Edgard Santos/Ebserh, Salvador, Bahia, Brazil

⁵ Infectious Disease Research Laboratory, Complexo Hospitalar Universitário Prof. Edgard Santos/Ebserh, Universidade Federal da Bahia, Salvador, Bahia, Brazil

Correspondence

Marcos A.D. Machado, Department of Nuclear Medicine, São Rafael Hospital/ Rede D'or, Salvador, Bahia 41253-190, Brazil.
Email: machado@radtec.com.br

Abstract

Purpose: Methodologies for optimization of SPECT image acquisition can be challenging due to imaging throughput, physiological bias, and patient comfort constraints. We evaluated a vendor-independent method for simulating lower count image acquisitions.

Methods: We developed an algorithm that recombines the ECG-gated raw data into reduced counting acquisitions. We then tested the algorithm to simulate reduction of counting statistics from phantom SPECT image acquisition, which was synchronized with an ECG simulator. The datasets were reconstructed with a resolution recovery algorithm and the summed stress score (SSS) was assessed by three readers (two experts and one automatic).

Results: The algorithm generated varying counting levels, simulating multiple examinations at the same time. The error between the expected and the simulated countings ranged from approximately 5% to 10% for the ungated simulations and 0% for the gated simulations.

Conclusions: The vendor-independent algorithm successfully generated lower counting statistics datasets from single-gated SPECT raw data. This method can be readily implemented for optimal SPECT research aiming to lower the injected activity and/ or to shorten the acquisition time.

KEYWORDS

gated SPECT, low-dose, optimization, raw data, shorter-scan-time

1 | INTRODUCTION

Single photon emission computed tomography (SPECT) is a widely used imaging method for the diagnosis and management of many diseases.^{1,2} The method consists of injecting radiopharmaceuticals to map their distribution in an organ. Optimizing the imaging protocols to lower the amount of administered radiopharmaceutical activities to patients, and/ or to shorten the acquisition duration, is recommendable.^{3–10}

Advances in detectors, reconstruction software, and imaging protocols allow for the reduction of counting statistics (which is a function of the administered activity and the acquisition duration) without compromising the image quality.^{4–17} However, defining the optimal counts requires to access various counting acquisition levels of the same patient.

Experimental schemes in image optimization studies play a key role in SPECT research. Physiological changes in patients between experiments, facility

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Journal of Applied Clinical Medical Physics* published by Wiley Periodicals, LLC on behalf of The American Association of Physicists in Medicine

imaging throughput, and patient comfort are important research constraints. Most studies scanned the patient two times using a variety of acquisition durations^{4–8} or preferably used list-mode acquisition.^{9–13} List-mode enables the use of the same image dataset to simulate various scanning durations and/or injected activity. With SPECT, only a few systems allow for simulating multiple scan durations^{9,10} or setting up multiple acquisition times concurrently.^{17–18} Therefore, vendor-independent methods that allow for nuclear medicine practitioners to easily simulate varying counting levels could facilitate the execution of protocol optimization studies.^{14–16}

The gated SPECT (G-SPECT) has been used in most clinical settings to assess the functional information (e.g., ejection fraction, wall motion, and thickness in cardiac studies).¹⁹ Since G-SPECT employs a partitioning method that splits the acquisition data into subsets of defined number of frames, it could be a promising candidate to make lower counts simulations widely available.¹⁴

In this study, we used G-SPECT to present a vendor-independent method to simulate varying amounts of counting levels.

2 | METHODS

In a typical ECG-gated SPECT acquisition, the cardiac cycle is divided into 8 or 16 frames, with each frame corresponding to a specific phase of the cardiac cycle in the tomographic projection.²⁰ The images acquired from the same frame are summated at each projection and subsequently reconstructed using all projections to generate the diagnostic image.

Our method consists of partitioning the cardiac cycle into η frames using the gamma-camera workstation, and then recombining a subset of these frames to a required counting level using our software, forming a new image with eight frames. Figure 1 illustrates a cardiac cycle divided into $\eta = 32$ frames. Following this, new simulated images can be reduced to 3/4 (Figure 1b), 1/2 (Figure 1c), 1/4 (Figure 1d), or recombined to the original ungated and gated images (Figure 1a).

An algorithm for implementation of the example in Figure 1e was developed in *Python*²¹ (Supporting information 1). After input of the G-SPECT image with η frames, the algorithm generates new datasets with simulated 100%, 75%, 50%, and 25% countings of the ungated and the gated images.

Phantom image acquisition and reconstruction were used to analyze our algorithm. According to Table 1, image acquisitions were performed with three Siemens SPECT systems using either the static cardiac phantom ECT/TOR/P (Data Spectrum Corporation)²² or the PET/SPECT Phantom Source Tank (76-823) with the Cardiac Insert (76-825) (Fluke Biomedical). The image acquisitions were performed synchronized either with an ECG simulator (Cardiac Trigger Monitor 7600) or syn-

TABLE 1 Experimental schemes

Scheme	SPECT system	ECG simulation	Cardiac phantom
1	e.cam	Cardiac Trigger Monitor 7600	ECT/TOR/P
2	Symbia Evo	Cardiac Trigger Monitor 7600	76-823 / 76–825
3	Symbia Intevo	Healthy volunteer	ECT/TOR/P

TABLE 2 Absolute error %Diff

Scheme	Simulation	A	B	C	D
1	gated	0.00%	0.00%	0.00%	0.00%
1	ungated	4.96%	4.98%	5.02%	4.95%
2	gated	0.00%	0.00%	0.00%	0.00%
2	ungated	9.65%	9.66%	9.64%	9.68%
3	gated	0.00%	0.00%	0.00%	0.00%
3	ungated	9.60%	9.60%	9.54%	9.57%

Note: A–D: 100%, 75%, 50%, and 25% simulated counts, respectively.

chronized with the heart cycle of a healthy volunteer (C.Q.) ($\eta = 32$ frames). Detailed imaging parameters are available at supporting information 2.

The image raw data were retrieved from the Siemens workstations and transferred to another workstation with the software, recombined into four datasets and then sent back to the Siemens workstation.

For each simulated image, the absolute error in the new datasets was estimated by calculating the percentage difference in the expected counting statistics, according to Equation (1):

$$\%Diff = 100 \cdot \left(\frac{E - S}{E} \right), \quad (1)$$

where E corresponds to the theoretical expected counting statistics and S corresponds to the counting statistics of 64 summed projections in the simulated image.

Projections were reconstructed using the Flash3D™ resolution recovery algorithm and the summed stress score (SSS) (17-segment polar map) was computed using Corridor4DM™ software (Michigan, United States).²³ Each simulated dataset was scored manually by two experienced nuclear medicine physicians, and scored automatically with the Corridor4DM™. The mean and 95% confidence interval (CI) of SSS was assessed for each counting level.

3 | RESULTS

The absolute error between the expected and the simulated counting statistics (Equation (1)) for schemes 1–3 is presented in Table 2.

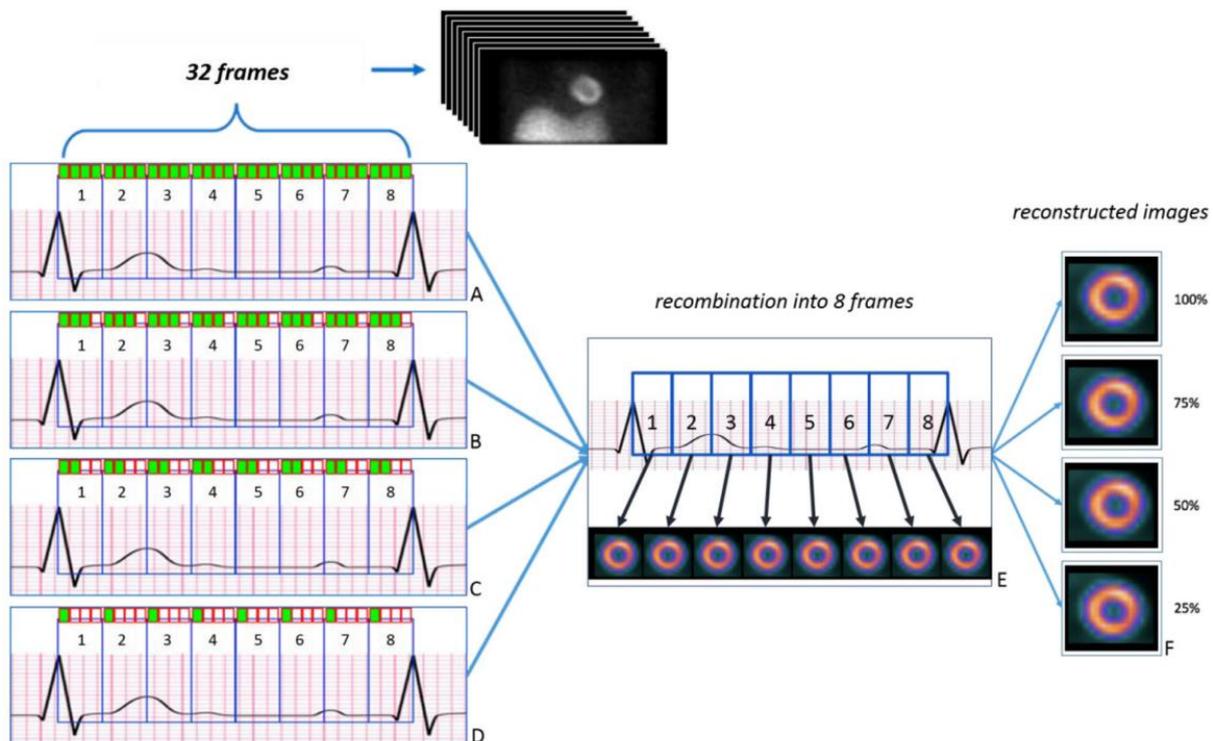


FIGURE 1 Partitioning and recombination of frames into lower counting statistics data. The colored marks indicate the selected frame to be rebinned into new simulated images. (A) 100% simulated statistics of the original data, 32 frames divided into 8 frames. (B) 75% simulated statistics of the original data, 24 frames divided into 8 frames. (C) 50% simulated statistics of the original data, 16 frames divided into 8 frames. (D) 25% simulated statistics of the original data, 8 frames divided into 8 frames. (E) Simulated dataset with reduced frames merged into its respective spot. (F) Reconstructed image of the summated frames

G-SPECT generates a standard ungated perfusion image by summing temporal frames together.²⁴ The Siemens systems, however, generate two independent gated and ungated datasets with 5%–10% less counts in the gated dataset. Since our software used the gated image to generate the simulated ungated, the reduced counts in the original gated image was propagated to the simulated ungated image.

The left-hand side of Figure 2 shows a projection of the partitioned/recombined raw data for the simulated ungated images. On the right is the reconstructed image showing the short, horizontal, and vertical axis of the phantom. Counting levels relative to the original input are shown in percentage terms as 100, 75, 50, and 25 from top to bottom. Figure 3 shows the simulated data for the gated image with 100% and 25% countings.

Figure 4 presents the polar map with the SSS quantifications for the reconstructed ungated images. The mean SSS value over three observers for 100%, 75%, 50%, and 25% counting levels were, respectively, 6.0 ± 2.0 , 8.3 ± 1.1 , 7.3 ± 1.1 and 7.3 ± 3.0 for the e.cam system; 0.3 ± 1.1 , 0.3 ± 1.1 , 0.0 ± 0.0 and 0.3 ± 1.1 for the Symbia Evo; and 9.0 ± 2.0 , 8.7 ± 3.0 , 8.3 ± 3.0 , and 8.7 ± 1.1 for the Symbia Intevo.

4 | DISCUSSION

We developed and provided a vendor-independent algorithm for raw data partitioning/ recombination of G-SPECT into varying counting levels to adequately simulate lower injected activity (or shorter acquisition duration) of SPECT studies. The feasibility of our method was validated in a phantom experiment and successfully produced reliable results.

Assessment of image quality through clinical indices has been widely reported.^{6–12,22} We showed a typical example of an optimization workflow to demonstrate the validity of our method. Experimental schemes resulted %Diff between 5% and 10%, suggesting older generation e.cam system yields slightly less counting loss in the gated image than current generation Symbia systems.

As expected, the reduction of the counting statistics did not significantly influence the SSS quantification. Lacchi and colleagues obtained similar trends at the same counting levels in a human model, where the SSS quantification was affected only for overweight and obese subjects.¹³ Here, we used a phantom that did not mimic the photon attenuation. Furthermore, we used a resolution recovery reconstruction algorithm, which is known to produce good image quality performance in

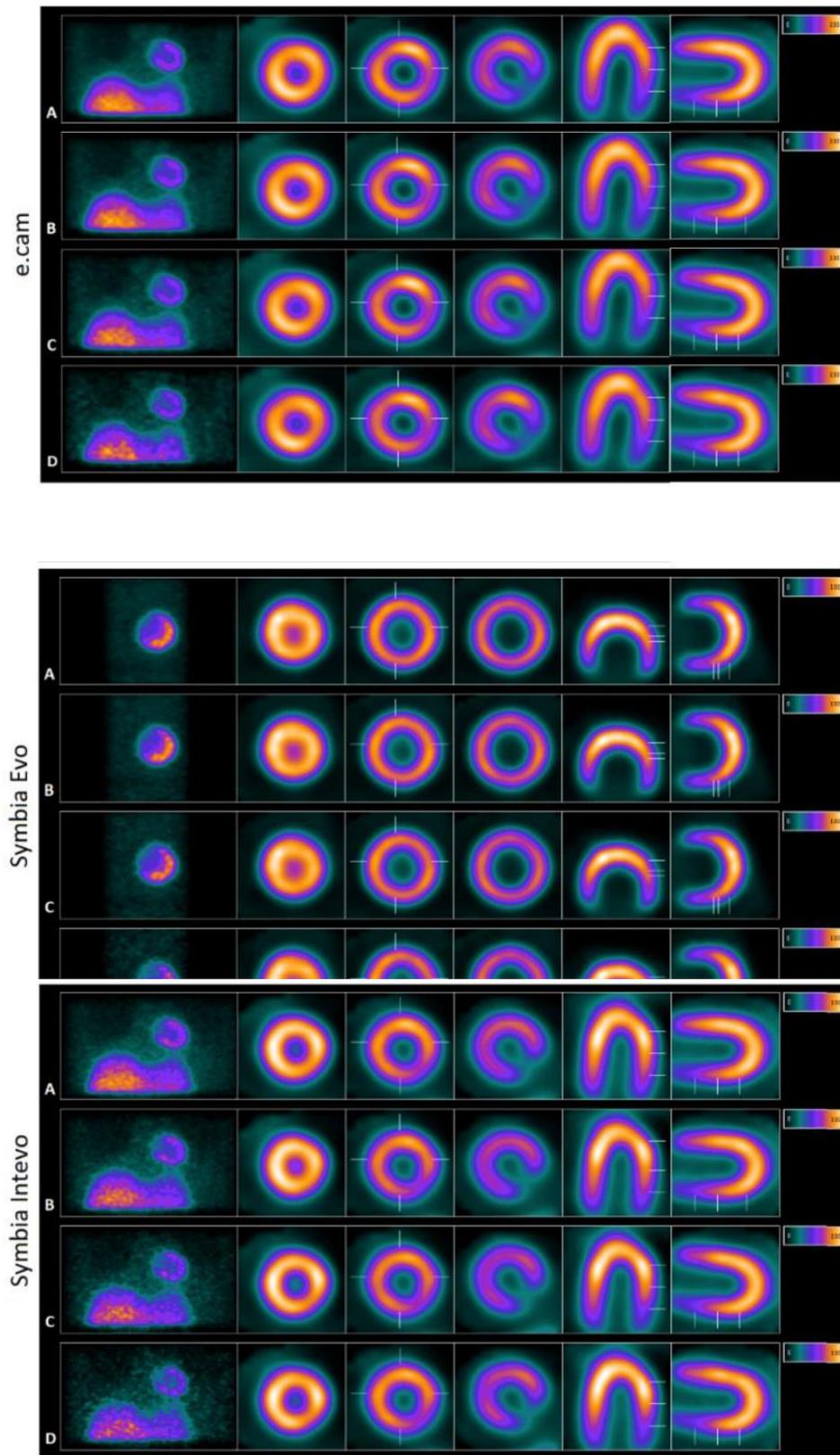


FIGURE 2 Simulated ungated projections (left) and reconstructed images (right). Simulated ungated imagens with 100% (A), 75% (B), 50% (C), and 25% (D) of the original raw data

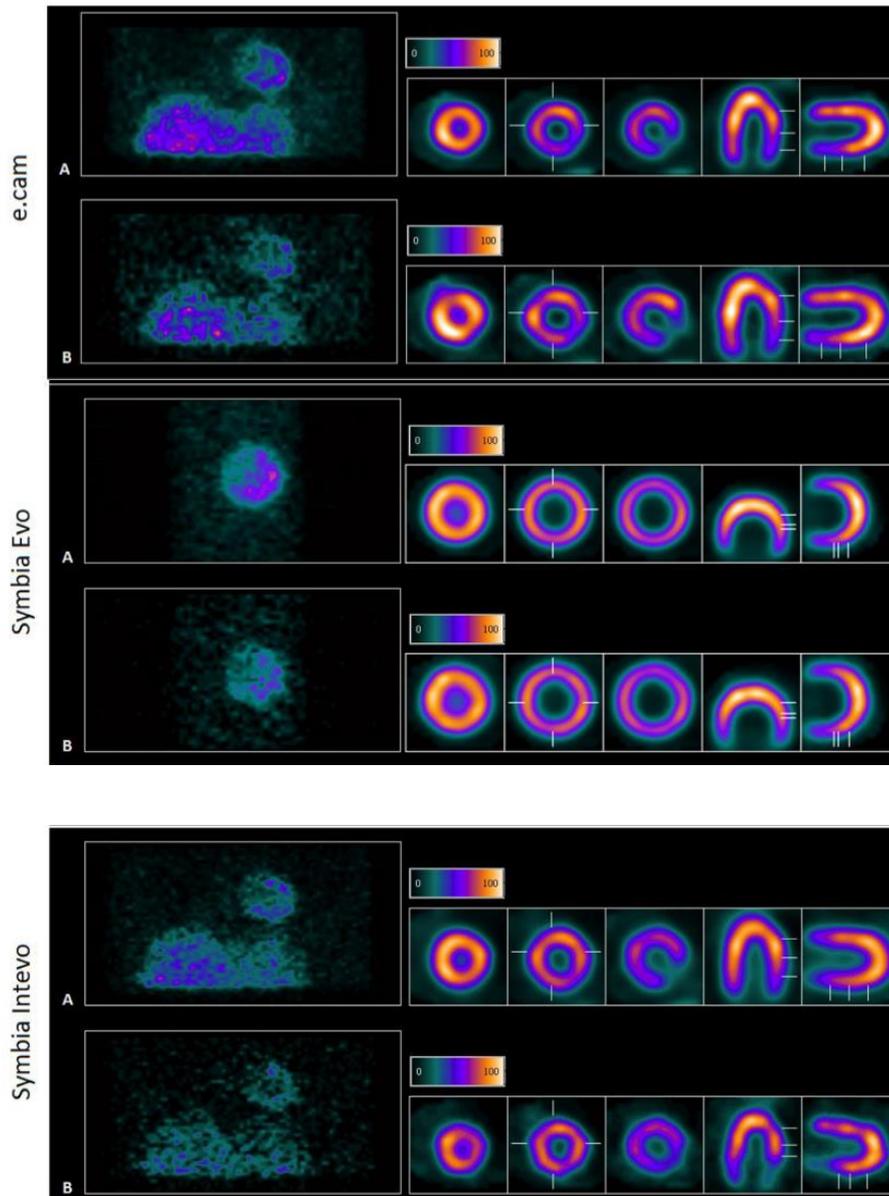


FIGURE 3 Gated projections (left) and reconstructed (right) images. Simulated statistics of the 8-frames gated data with 100% (A) and 25% (B) counts of the original raw data

various count levels.^{4-8,11,13,22} In addition, the phantom model produces images from the static activity distributions, while a real patient model would have moving organs and thus lower quality images. Then, these may

In this study, we used $\eta = 32$ frames which is the maximum setup available at the Siemens systems. Rebinning η into 8 frames using our software results in a standard cardiac-gated image and enables the recovery of perfusion and gated information. This application of our method is appropriate for the most complex human model, where capturing the information (ECG-synchronized) of the moving organ is required. Moreover, any simpler rebinnings can provide lower

counting statistics for other clinical applications such as neurology and oncology.⁴ This example is also available (supporting information 1) to generate new datasets with $1/\eta$ rebinnings.

We therefore provided a vendor-independent algorithm which is readily applicable to different SPECT exams. This technique is easy to implement, able to be performed rapidly, and is highly robust. Our method improves the image optimization process since the patient is submitted to one only image acquisition, which increases patient comfort, while enhancing the routine throughput during optimization studies.

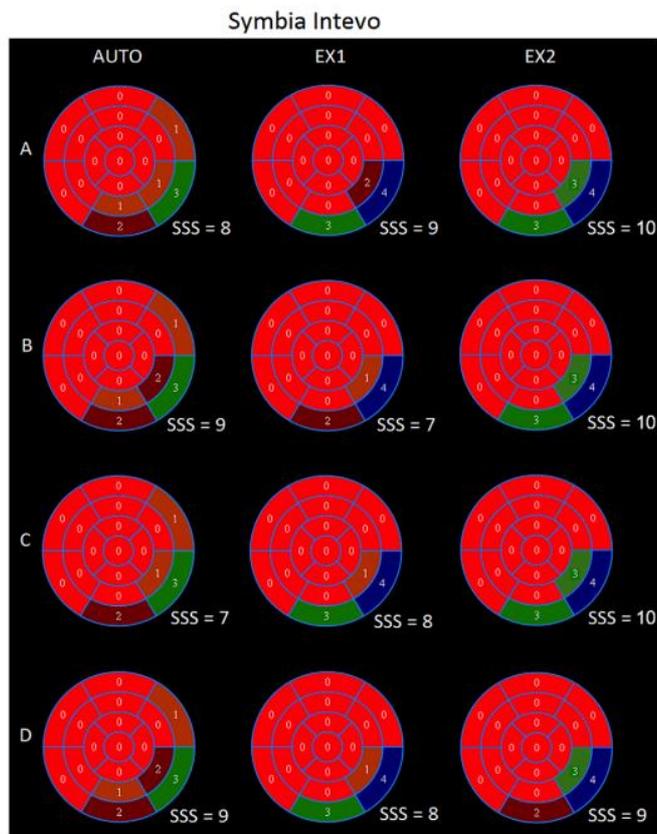
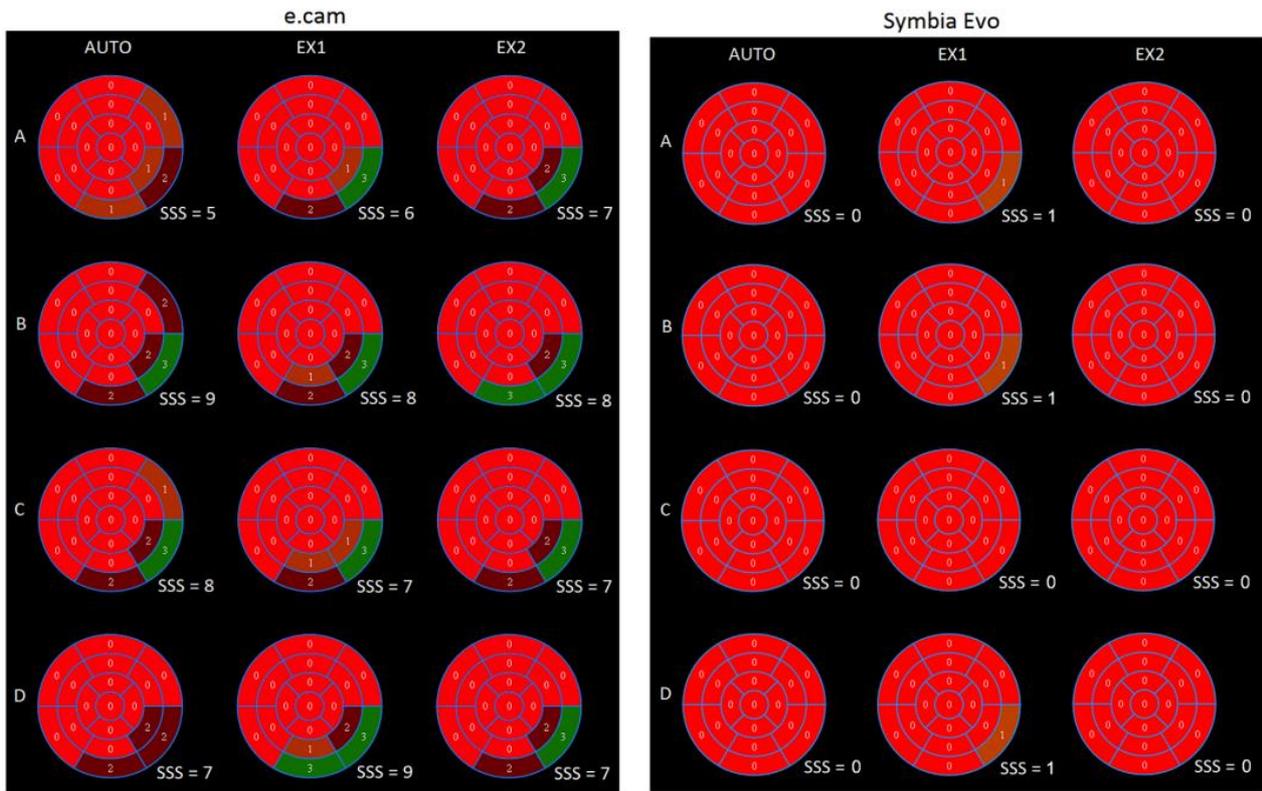


FIGURE 4 polar map quantifications for the simulated ungated images. (A–D) represents quantifications of the ungated simulations for 100%, 75%, 50%, and 25% of the original data statistics, respectively. AUTO: automatic quantification using standard database of normality. EX1 and EX2: nuclear medicine expert quantifications. SSS: summed stress score for 17-segment polar map

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

ACKNOWLEDGMENTS

We appreciate the support of Programa de Pós-Graduação em Medicina e Saúde, Faculdade de Medicina da Bahia, Universidade Federal da Bahia, Salvador, Bahia, Brazil; and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brazil (CAPES)-Finance Code 001. We also thank Bruno Cesar da Rocha Farias Santana for his contribution with the experiments and data collection.

AUTHOR CONTRIBUTIONS

Cleiton C. Queiroz contributed to study design, data collection, algorithm development, data analysis, interpretation of results, and final approval of the version to be published. Marcos A. D. Machado contributed to study conception and design, data analysis and interpretation of results, drafting the article, and final approval of the version to be published. Antonio A. B. Ximenes contributed to data analysis, revising the article, and final approval of the version to be published. Andre G. S. Pino contributed to data analysis, revising the article, and final approval of the version to be published. Eduardo M. Netto contributed to interpretation of results, revising the article, and final approval of the version to be published.

REFERENCES

- Ritt P, Vija H, Hornegger J, Kuwert T. Absolute quantification in SPECT. *Eur J Nucl Med Mol Imaging*. 2011;38(Suppl 1):S69-S77.
- Mariani G, Bruselli L, Kuwert T, et al. A review on the clinical uses of SPECT/CT. *Eur J Nucl Med Mol Imaging*. 2010;37:1959-1985.
- Menezes VO, Machado MAD, Queiroz CC, et al. How does Lean Six Sigma method improve healthcare practice in nuclear medicine departments? A successful case of dedicated software applications in oncological PET/CT. *J Nucl Med*. 2018;59(Suppl 1):1013.
- Zacho HD, Manresa JAB, Aleksyniene R, et al. Three-minute SPECT/CT is sufficient for the assessment of bone metastasis as add-on to planar bone scintigraphy: prospective head-to-head comparison to 11-min SPECT/CT. *EJNMMI Res*. 2017;7(1):1-7.
- Valenta I, Treyer V, Husmann L, et al. New reconstruction algorithm allows shortened acquisition time for myocardial perfusion SPECT. *Eur J Nucl Med Mol Imaging*. 2010;37(4):750-757.
- Borges-Neto S, Pagnanelli RA, Shaw LK, et al. Clinical results of a novel wide beam reconstruction method for shortening scan time of Tc-99m cardiac SPECT perfusion studies. *J Nuc Cardiol*. 2007;14(4 SPEC. ISS.):555-565.
- DePuey EG, Gadiraju R, Clark J, Thompson L, Anstett F, Shwartz SC. Ordered subset expectation maximization and wide beam reconstruction "half-time" gated myocardial perfusion SPECT functional imaging: a comparison to "full-time" filtered backprojection. *J Nuc Cardiol*. 2008;15(4):547-563.
- Druz RS, Phillips LM, Chugkowski M, Boutis L, Rutkin B, Katz S. Wide-beam reconstruction half-time SPECT improves diagnostic certainty and preserves normalcy and accuracy: a quantitative perfusion analysis. *J Nuc Cardiol*. 2011;18(1):52-61.
- Nakazato R, Berman DS, Hayes SW, et al. Myocardial perfusion imaging with a solid-state camera: simulation of a very low dose imaging protocol. *J Nuc Med*. 2013;54(3):373-379.
- Palyo RJ, Sinusas AJ, Liu YH. High-sensitivity and high-resolution SPECT/CT systems provide substantial dose reduction without compromising quantitative precision for assessment of myocardial perfusion and function. *J Nuc Med*. 2016;57(6):893-899.
- Machado MAD, Menezes VO, Namias M, et al. Protocols for harmonized quantification and noise reduction in low-dose oncologic 18F-FDG PET/CT imaging. *J Nuc Med Technol*. 2019;47(1):47-54.
- Zhang J, Liu X, Knopp MI, Ramaswamy B, Knopp MV. How Long of a Dynamic 3'-Deoxy-3'-[18 F] fluorothymidine ([18 F] FLT) PET acquisition is needed for robust kinetic analysis in breast cancer? *Mol Imaging Biol*. 2019;21(2):382-390.
- Menezes VO, Machado MAD, Queiroz CC, et al. Optimization of oncological 18F-FDG PET/CT imaging based on a multiparameter analysis. *Med Phys*. 2016;43(2):930-938. <https://aapm.onlinelibrary.wiley.com/doi/10.1118/1.4940354>
- Armstrong IS, Arumugam P, James JM, Tonge CM, Lawson RS. Reduced-count myocardial perfusion SPECT with resolution recovery. *Nucl Med Commun*. 2012;33(2):121-129.
- Kim IH, Lee SJ, An YS, Choi SY, Yoon JK. Simulating dose reduction for myocardial perfusion SPECT using a Poisson resampling method. *Nucl Med Mol Imaging*. 2021;55(5):245-252. <https://doi.org/10.1007/s13139-021-00710-w>
- Lawson RS, White D, Nijran K, et al. An audit of half-count myocardial perfusion imaging using resolution recovery software. *Nucl Med Commun*. 2014;35(5):511-521.
- Lecchi M, Martinelli I, Zoccarato O, Maioli C, Lucignani G, Del Sole A. Comparative analysis of full-time, half-time, and quarter-time myocardial ECG-gated SPECT quantification in normal-weight and overweight patients. *J Nucl Cardiol*. 2017;24(3):876-887.
- Coles DE, Murray D, Bertelsen H. Multiple simultaneous acquisition of gamma camera data sets. US patent 6,900,441. May 31, 2005.
- Paul AK, Nabi HA. Gated myocardial perfusion SPECT: basic principles, technical aspects, and clinical applications. *J Nucl Med Technol*. 2004;32(4):179-187.
- Montelatici G, Sciagrà R, Passeri A, et al. Is 16-frame really superior to 8-frame gated SPECT for the assessment of left ventricular volumes and ejection fraction? Comparison of two simultaneously acquired gated SPECT studies. *Eur J Nucl Med Mol Imaging*. 2008;35:2059-2065.
- Van Rossum G, Drake FL. *Python 3 Reference Manual*. CreateSpace; 2009.
- Zoccarato O, Scabbio C, Ponti E, et al. Comparative analysis of iterative reconstruction algorithms with resolution recovery for cardiac SPECT studies. A multi-center phantom study. *J Nucl Cardiol*. 2014:135-148.
- Ficaro EP, Lee BC, Kritzman JN, et al. Corridor4DM: the Michigan method for quantitative nuclear cardiology. *J Nucl Cardiol*. 2007;14:455-465.
- Paul AK, Nabi HA. Gated myocardial perfusion SPECT: basic principles, technical aspects, and clinical applications. *J Nucl Med Technol*. 2004;32(4):179-187.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Queiroz CC, Machado MAD, Ximenes AAB, Pino AGS, Netto EM. Technical note: Partitioning of gated single photon emission computed tomography raw data for protocols optimization. *J Appl Clin Med Phys*. 2022;23:e13508. <https://doi.org/10.1002/acm2.13508>

Artigo nº 2

**PET Radiomic Features Variability: A Phantom Study on the Influence of
Reconstruction and Discretization Method**

Revista Brasileira de Física Médica

Publicado

PET Radiomic Features Variability: A Phantom Study on the Influence of Reconstruction and Discretization Method

Variabilidade de Atributos Radiômicos em PET: Um Estudo com Simuladores sobre a Influência dos Métodos de Reconstrução e Discretização

Marcos A. D. Machado^{1,2}, Jean L. P. Pita³, Eduardo M. Netto²

¹São Rafael Hospital, Nuclear Medicine Department, Salvador, Brazil

²Complexo Hospitalar Universitário Professor Edgard Santos/UFBA/Ebserh, Salvador, Brazil

³UFBA/Computer Science Department, Salvador, Brazil

Abstract

We aimed to explore the variability of PET radiomic features for varying reconstruction methods and quantification settings. The IQ-NEMA phantom was scanned five times with a sphere to background F-18 concentration ratio of 10:1. The activity and the scan duration were matched to result in typical counting statistics for 18F-FDG oncologic examinations. The images were reconstructed with OSEM and PSF reconstructions, then 99 radiomic features were extracted using two discretization methods: fixed bin number (FBN = 16, 32 and 64 gray levels) and fixed bin width (FBW=0.25). This scheme resulted in a total of 1,188 features, classified as having low (<5.0%), intermediate (5-29.9%) or high (≥30%) variability. In general, FBW discretization yielded more stable features. A total of 499, 558 and 131 features had low, intermediate and high variability, respectively. First-order features such as energy and entropy and textural features such as entropy (GLCM), long run emphasis and short run emphasis (GLRLM) were more likely to present low variability, regardless the reconstruction and discretization method. Other textural features such as large area emphasis (GLSZM), zone percentage (GLSZM) and complexity (NGTDM) had more frequently intermediate or high variability. These findings could facilitate features' selection for further PET radiomic applications.

Keywords: PET radiomic features; variability; quantification; standardization.

Resumo

O objetivo deste estudo é explorar a variabilidade de atributos radiômicos do PET para diferentes métodos de reconstrução e parâmetros de quantificação. Foram realizadas cinco imagens do simulador IQ-NEMA com uma razão de concentração esfera-fundo de F-18 de 10:1. A atividade radiativa e o tempo de aquisição das imagens foram combinados para alcançar uma estatística de contagens típica de exames oncológicos com 18F-FDG. As imagens foram reconstruídas com os métodos OSEM e PSF, obtendo-se em seguida a quantificação de 99 marcadores de imagens a partir de dois métodos de discretização: fixed bin number (FBN = 16, 32 e 64 níveis de cinza) e fixed bin width (FBW=0.25). Esta configuração resultou em 1.188 elementos de imagens (radiomic features), classificados quanto à variabilidade como baixa (<5.0%), intermediária (5-29.9%) ou alta (≥30%). Em geral, o método FBW resultou em atributos mais estáveis. Um total de 499, 558 e 131 atributos tiveram variabilidades classificadas como baixa, intermediária e alta, respectivamente. Atributos de primeira ordem como energia e entropia, e marcadores de textura como entropia (GLCM), long run emphasis e short run emphasis (GLRLM) foram classificados como baixa variabilidade independentemente dos métodos de reconstrução e quantização. Outros atributos de textura como large area emphasis (GLSZM), zone percentage (GLSZM) e complexity (NGTDM) foram classificados mais frequentemente como variabilidade intermediária ou alta. Estes achados podem facilitar a escolha e seleção de atributos de imagens em aplicações futuras com radiômica em PET.

Palavras-chave: Características Radiômicas do PET; variabilidade; quantificação; padronização.

1. Introduction

Positron emission tomography (PET) and computed tomography (CT) hybrid imaging (PET/CT) is widely used for clinical diagnosis, staging, prognosis and treatment response assessment (1,2). The ability to use imaging biomarkers such as the standardized uptake value (SUV) and metabolic tumor volume (MTV) emphasizes this type of PET role (3,4). It is a consensus today that the novel PET biomarkers based on texture, morphological and metabolic features (radiomics) provide substantially more information about the disease than the current imaging practices (5–11). In this regard, there is a growing interest to assess PET radiomics features stability and reproducibility (8,12,13).

Radiomic features are affected by several factors, including image acquisition and reconstruction

parameters, the feature calculation workflow as the image segmentation, discretization, feature mathematical design and calculation settings. However, the lack of standardisation in PET radiomic research is limiting its potential new applications (9)(14). Some authors have studied different robustness aspects to identify potential radiomic features or clusters (13,15,16). Nyflot et al (16) used realistic phantom simulations to evaluate the effects of noise and patient size on the feature variability (type II error) and image reconstruction on metrics bias (type I error). They used a fixed feature calculation setting of 256 gray level discretization and estimated the sample size that would be required for clinical trials to power for clinical effects of 30% and 15% due to only stochastic variability. They found a large and feature-dependent PET features sensitivity

to imaging parameters. Pfaehler et al.¹³ on the other hand, explored PET radiomic feature variability using IEC/NEMA and 3D printed phantoms. They used different image reconstruction, noise and segmentation methods on the composition of feature clusters and studied two discretization methods: fixed bin number (FBN) of 64 gray levels and fixed bin width (FBW) of 0.25 and 0.05. Their results confirmed the sensitivity of PET radiomic features to imaging parameters and suggested image acquisition and processing standardization.

Despite the increased use of radiomic for PET imaging, relatively little is known about the image reconstruction and quantification parameters impact on feature variability, limiting its standardization and clinical applications. Understanding the image biomarkers' quantitative aspects may simplify the features' selection, so this study uses IEC/NEMA phantom experiments to analyze PET radiomic features variability with respect to reconstruction methods and discretization parameters.

2. Materials and Methods

2.1. Scanner characteristics

PET/CT imaging was performed on a LSO-based PET Siemens Biograph TruePoint TrueV (Knoxville, TN, USA) combined with a 16-slice helical CT scanner (Emotion 16; Siemens). PET images were corrected for random coincidences, normalization, dead time losses, scatter and attenuation. The attenuation map was obtained by a spiral CT scan (100 kVp, automatic tube-current modulation), 3 mm slice thickness and a standard soft tissue reconstruction kernel (Siemens B30s).

2.2. Phantom preparation and imaging

We used an IEC/NEMA body phantom with 6 spheres (internal diameters of 10, 13, 17, 22, 28 and 37 mm) with a sphere to background F-18 concentration ratio of 10:1, according to the EANM guideline (17). To evaluate feature variability, five PET/CT images were sequentially acquired with two bed positions and the activity and the scan duration matched to result 416 MBq.s/kg (9.4 kg was the phantom weight considered in background concentration). This counting statistic represents typical whole-body oncologic examination (17).

2.3. Phantom image reconstruction

Images were reconstructed on a 168x168 matrix size (4.07 x 4.07 mm² voxels) with 3 mm slice thicknesses. Three reconstruction algorithms were used: (1) Ordered-Subsets Expectation-Maximization with 3 iterations, 21 subsets and 5 mm Gaussian filter (OSEM3D) (EARL-compliant); (2) Point Spread Function Ordinary Poisson with 3 iterations, 21 subsets and 7 mm Gaussian filter (PSF7) (EARL-compliant); and (3) Point Spread Function Ordinary Poisson with 2 iterations, 21 subsets and 2 mm Gaussian filter (PSF2). Both reconstructions PSF7 and OSEM3D are compliant with the EANM standards for quantification, being the PSF2 the option of choice for detectability purposes (18).

2.4. Radiomic feature extraction

Spheres were segmented with the Beth Israel PET/CT plugin for FIJI (ImageJ, Bethesda, MD, USA)¹⁸ using the automatic segmentation with 41% threshold of maximum SUV (SUVmax). The segmented volumes of interest (VOIs) were analyzed with PyRadiomics, an open-source platform available at www.radiomics.io that enables a large panel of radiomic features extraction (21). We resampled the matrix grid to cubic voxels of 4x4x4 mm³ using B-Spline interpolation and the discretization within each VOI were scaled to FBN = 16, 32 and 64 grayscale levels and FBW = 0.25. The radiomic feature classes and corresponding features as defined by Griethuysen et al (21) are presented in the supplementary material. A total of 99 features were calculated, including 18 first order and 81 textural features: 16 gray level run length matrix (GLRLM), 16 gray level size zone matrix (GLSZM), 5 neighborhood gray-tone difference matrix (NGTDM) and 22 gray level co-occurrence matrix (GLCM). GLCM features were computed by 2 different methods: using 13 matrices, one calculated for each spatial direction separately, after which the mean of these values is returned (method A); and using only one matrix, weighted by factor 1 and summed, without the average step (method B) (22).

2.5. Variability estimation and analysis

Feature variability was estimated through a variation of the metric used for maximum Standard Uptake Value (SUVmax) (18) by calculating the average coefficient of variation for each reconstruction and discretization as follows:

$$Q_{var} = \frac{\sum_{i=4}^6 (\sigma_i/M)}{3} \quad (1)$$

Where σ_i is the standard deviation of the feature across realizations for sphere i , and M is the mean value of the feature. The 10 mm, 13 mm and 17 mm spheres were excluded from the analysis due to insufficient pixels for quantification of textural features. Only the three largest spheres were included (volumes: 5.6, 11.5 and 26.5 cm³).

Features were classified into three ranges of variability: low (<5.0%), intermediate (5-29.9%) and high ($\geq 30.0\%$) and the total number of stable features was then used as a measure to investigate the influence of reconstruction and discretization on feature stability.

Low variability level was defined based on our previous efforts to study the variability of SUVmax (18), while the intermediate and high levels were defined based on the rationale presented by Galavis et al (15).

3. Results

A total of 499, 558 and 131 features quantifications were ranked as low, intermediate, and high, respectively (Table 1). PSF7 (FBN=32) was the reconstruction that yielded more features with low variability (n=59); and PSF7 (FBN=16) was those with

higher variability (n=55). PSF7 was, therefore, very sensitive to the number of bins. Concerning the discretization method, FBW=0.25 and FBN=32 presented the smallest number of features with high variability (n=14).

Table 1 - Number of features according to variability classification.

	FBW:0.25	FBN:16	FBN:32	FBN:64
Low				
OSEM	56	46	45	47
PSF2	37	32	37	34
PSF7	36	20	59	50
Intermediate				
OSEM	38	51	51	41
PSF2	57	61	53	56
PSF7	59	24	38	29
High				
OSEM	5	2	3	11
PSF2	5	6	9	9
PSF7	4	55	2	20

Variability: low (<5.0%), intermediate (5-29.9%) and high (≥30.0%). FBN: fixed bin number. FBW: fixed bin width.

Figure 1 presents the variability (Qvar) of 20 selected features for improved readability, where GLCM is presented for the setup using only one array considering all 13 directions simultaneously, without an average step. The features in Figure 1 were selected based on previous reports (10,13,15,16,22,23), the complete list of all 99 features is available in the supplementary material (Table S2). The feature named Dissimilarity reported in other studies (10,16) has been discontinued in Pyradiomics, as it is mathematically equal to Difference Average (GLCM) (21); and high intensity large area emphasis (HILAE - GLSZM) reported by Hatt et al²¹ is mathematically equal to large area emphasis, the nomenclature used in Pyradiomics (Figure 1). Colour blue, yellow, and red highlights the features classified as low, intermediate, and high variability, respectively. Some quantifications were possibly negatively biased because outliers, since the standard deviation in equation 1 is sensitive to replications number and hence the Qvar. For example, the feature Large Area Emphasis (GLSZM) at OSEM reconstruction (FBN=64) achieved high variability (Qvar = 80.7%) while its quantification at FBW = 0.25 had an intermediate variability (Qvar = 14.2%). The higher variability is solely explained by

one measurement three-fold higher in one of the spheres.

The presented variability in Table 1 and Table S2 support future radiomic studies by selecting the adequate reconstruction, radiomic features and discretization method. For example, for a given reconstruction algorithm (eg.: OSEM or PSF) and discretization method (eg.: FBN=64 or FBW=0.25), one may select a subset of features from Figure 1 (eg.: 10-20 features) to investigate their disease relationship.

4. Discussion

This study presents a pool of features that might be used to simplify future PET radiomic applications. We classified radiomic features according to their variability performance for different discretization methods and using reconstruction algorithms currently used in clinical practice (24–26). However, the variability analysis was performed through a simple metric and phantom preparation used to characterize the SUVmax variability with respect to image noise and reconstruction algorithm previously reported by Machado and colleagues (18). It was found a 1.5% < Qvar < 4.6% for SUVmax, a biomarker widely used in several oncologic PET studies.

We found also that the textural features presented worst variability more frequently than first-order features. Pfähler and colleagues (13) used 10 replications of the IEC/NEMA and 3D printed phantoms to demonstrate the impact of contrast (sphere to background concentration ratio), image reconstruction, noise, discretization and segmentation methods on the repeatability of PET radiomic features through intraclass correlation coefficients (ICC) and cluster analysis. Here, we evaluated additional discretization parameters and assessed variability through a simpler method. Despite the agreement of variability levels of several features (for matched reconstruction and discretization), the direct comparison with our results is challenging because of differences in metrics used (Qvar versus ICC - data shown for EARL-compliant reconstruction (13)) and the number of features (here, 99 features versus 246 features in their work, we did

Features	OSEM				PSF2				PSF7			
	FBN=64	FBN=32	FBN=16	FBW=0.25	FBN=64	FBN=32	FBN=16	FBW=0.25	FBN=64	FBN=32	FBN=16	FBW=0.25
first order_Energy	3.0%	3.0%	3.0%	2.8%	3.1%	3.1%	3.1%	2.6%	2.1%	2.1%	2.1%	2.0%
first order_Entropy	1.3%	1.3%	1.7%	1.3%	2.2%	2.6%	3.0%	1.5%	1.0%	0.9%	17.2%	2.0%
first order_Kurtosis	4.0%	4.0%	4.0%	3.0%	4.5%	4.5%	4.5%	6.2%	2.2%	2.2%	2.2%	2.3%
first order_Skewness	48.0%	48.0%	48.0%	20.0%	80.0%	80.0%	80.0%	20.0%	22.0%	22.0%	22.0%	7.0%
glcm_ClusterProminence	13.7%	14.0%	13.3%	9.1%	30.9%	31.0%	31.8%	15.8%	10.8%	10.5%	219.2%	24.1%
glcm_Correlation	7.0%	7.0%	6.8%	4.6%	23.4%	23.3%	22.7%	12.0%	6.3%	6.1%	6.1%	8.8%
glcm_DifferenceAverage	3.3%	3.3%	3.3%	2.4%	6.5%	6.5%	7.0%	3.9%	1.6%	1.6%	83.5%	3.7%
glcm_JointEntropy	1.0%	1.2%	1.5%	1.3%	1.2%	1.5%	2.3%	1.1%	0.9%	0.9%	6.8%	1.0%
glrlm_HighGrayLevelRunEmphasis	7.1%	6.7%	8.8%	6.8%	15.3%	13.7%	14.0%	12.5%	2.9%	3.2%	162.4%	13.3%
glrlm_LongRunEmphasis	0.8%	1.7%	3.0%	1.9%	1.1%	1.6%	2.8%	1.7%	1.0%	1.2%	7.9%	1.3%
glrlm_LongRunHighGrayLevelEmphasis	6.7%	5.8%	7.4%	6.3%	15.1%	13.3%	13.4%	12.9%	4.3%	5.1%	148.9%	13.1%
glrlm_ShortRunEmphasis	0.2%	0.5%	0.8%	0.5%	0.3%	0.5%	0.7%	0.5%	0.3%	0.3%	2.2%	0.4%
glszm_HighGrayLevelZoneEmphasis	81.2%	10.6%	11.2%	11.0%	15.2%	13.7%	11.7%	13.2%	3.3%	4.5%	171.3%	14.2%
glszm_LargeAreaEmphasis	80.7%	20.0%	15.5%	14.2%	14.6%	90.4%	23.4%	25.8%	153.0%	14.2%	46.4%	13.7%
glszm_SmallAreaEmphasis	1.9%	3.6%	6.6%	5.4%	2.3%	1.9%	4.8%	3.9%	2.0%	3.2%	15.3%	5.1%
glszm_ZonePercentage	2.9%	77.2%	6.7%	6.4%	76.9%	78.6%	81.0%	6.9%	3.5%	77.7%	30.3%	4.7%
ngtdm_Busyness	14.0%	11.4%	14.6%	7.5%	14.1%	11.9%	12.9%	13.2%	7.4%	9.2%	48.6%	18.4%
ngtdm_Coarseness	10.1%	9.2%	9.5%	5.2%	6.7%	7.1%	5.3%	4.8%	6.5%	7.1%	13.6%	6.6%
ngtdm_Complexity	8.1%	8.6%	10.8%	7.3%	18.9%	15.1%	15.6%	11.6%	4.2%	4.6%	202.3%	16.2%
ngtdm_Contrast	10.4%	8.9%	9.3%	8.9%	11.1%	15.0%	19.0%	13.2%	8.3%	8.8%	111.7%	7.5%

Figure 1 - Quantification variability of 20 selected features. Variability: low (<5.0%), intermediate (5-29.9%) and high (≥30.0%) are highlighted in colours blue, yellow and red, respectively.

not include in this work morphological features and 2D measures of textural features) included in the analysis. Nevertheless, our results are also in concordance with Pfaehler et al (13) that FBW=0.25 yields more stable features (Qvar low and intermediate) than FBN=64, but we found that FBN=32 were equivalent to FBW=0.25 respective to stable features number. In addition, resampling with FBN lower to 32 gray levels (eg.: FBN = 16) should be considered with caution due to the high correlation among features (10). Indeed, each feature shall be analyzed individually before choosing predictors of clinical outcomes. Figure 1 shows that some features are highly stable regardless the reconstruction and/or discretization method.

Pfaehler et al (13) also demonstrated that lower uptake regions are more sensitive to the segmentation method. Our study only used 41% threshold of SUVmax and used spheres with radiopharmaceutical concentrations 10:1 with respect to the background. In such case the radiopharmaceutical concentration is considerably higher than background, thus allowing for more consistently contouring the sphere, regardless the segmentation method. For this reason, our results may be extrapolated only to evaluations where lesions have radiopharmaceutical uptake concentration of 10:1 or higher, otherwise segmentation constraints shall be observed.

This study also evaluated two different methods for computation of GLCM features. A markedly improvement in variability was observed in PSF7 FBN=64 when method A was applied. In general, the features Cluster Prominence and Cluster Shade performed better in method B for PSF2 and FBN=16, 32 and 64; the feature Maximum Probability performed better in method A for all quantization settings; and the feature Joint Average significantly worsen in method B, PSF2, FBN=64.

Nyflot et al simulated 50 statistically independent IQ NEMA images from a GE D670 PET/CT scanner with PSF reconstruction (5 mm Gaussian filter) and performed quantifications with FBN = 256 (16). They also found complex trends in the coefficient of variation (variability) as a function of features, sphere, patient sizes and reconstruction parameters (by changing the iterations number from 2 to 6 in the reconstruction process).

It is important to stress that feature's choice also depends on the disease (22). For example, a given feature with intermediate variability may be clinically better than other with lower; since feature variation is sufficiently large to detect a difference in disease presentation. Because several features are highly correlated and redundant, a maximum number of around ten features is usually used at the reduced space preserving the information that explains a biological or clinical phenomena (10,22,23,27–29).

A heuristic comparison among our findings with those found in Pfaehler (13) and Nyflot et al (16), identified common features with reasonable variability, for example: 1) energy and entropy from first order statistics metrics, 2) difference average and

entropy from GLCM, 3) long run emphasis and short run emphasis from GLRLM, 4) zone percentage and large area emphasis from GLSZM, and 5) complexity and coarseness from NGTDM. Therefore, we recommend that these features are included in further PET radiomic studies to simplify the study design.

A limitation of this study is the use of only high uptake and homogenous spheres, not mimicking a real clinical condition. However, we presented a starting point where features, reconstruction method and quantification setup might be mined for further radiomic research. An additional drawback in radiomic studies are variations in nomenclature and formulas among studies and quantification software (9), which have recently been subjected to standardization along with other reporting standards (30). Here, we highlighted the features HILAE and Difference Average, which have been reported to yield clinical significance and presented low to intermediate variability in most reconstruction and quantization setups.

In clinical practice, the use of PSF reconstruction is preferable because it allows for improved lesion detectability and reduced image noise (18,25,31). Our results present several radiomic feature variability measures using two PSF reconstruction settings. In addition, the ability of PSF reconstructions to improve the system spatial resolution is also expected to detect smaller differences in textural features. This study did not assess type I error, which informs the feature variation between different scanners. Thus, one must be careful when selecting radiomic features for multicenter studies and harmonization techniques shall be considered (32).

The radiomic features extraction provides a powerful method to assess differences in tumor biology by identifying predictors that may have a functional role in specific phenotypes. Our results contribute to overcoming standardization challenges that need to be addressed before radiomics can safely be implemented in the clinic.

5. Conclusions

Our data show the PET radiomic features' sensitivity regarding the reconstruction and discretization method, allowing researchers to assess a starting point for feature selection, reconstruction and discretization's choice in further PET radiomic studies. Some often reported features were likely low variability, regardless of the reconstruction and discretization method, while others (mostly textural features) were likely intermediate or high variability depending on the reconstruction and/ or discretization setup.

Acknowledgements

Programa de Pós-Graduação em Medicina e Saúde, Faculdade de Medicina da Bahia, Universidade Federal da Bahia, Salvador, Bahia, Brazil.

References

1. SCHREVEVS L, LORENT N, DOOMS C, VANSTEENKISTE J. The Role of PET Scan in Diagnosis, Staging, and Management of Non-Small Cell Lung Cancer. *Oncologist*. 2004;9:633–43.
2. Barrington SF, Mikhaeel NG, Kostakoglu L, Meignan M, Hutchings M, Müller SP, et al. Role of imaging in the staging and response assessment of lymphoma: Consensus of the international conference on malignant lymphomas imaging working group. *J Clin Oncol*. 2014;32(27):3048–58.
3. Wahl RL, Jacene H, Kasamon Y, Lodge MA. Response Criteria in Solid Tumors. *J Nucl Med*. 2009;50(Suppl 1):122–50.
4. Kanoun S, Rossi C, Casasnovas O. [18F]FDG-PET/CT in hodgkin lymphoma: Current usefulness and perspectives. *Cancers (Basel)*. 2018;10(5):1–11.
5. Cook GJR, Yip C, Siddique M, Goh V, Chicklore S, Roy A, et al. Are Pretreatment 18F-FDG PET Tumor Textural Features in Non-Small Cell Lung Cancer Associated with Response and Survival After Chemoradiotherapy? *J Nucl Med*. 2013;54(1):19–26.
6. Cook GJR, Siddique M, Taylor BP, Yip C, Chicklore S, Goh V. Radiomics in PET: Principles and applications. *Clin Transl Imaging*. 2014;2(3):269–76.
7. De Geus-Oei LF, Van Der Heijden HFM, Corstens FHM, Oyen WJG. Predictive and prognostic value of FDG-PET in non-small-cell lung cancer. A systematic review. *Cancer*. 2007;110(8):1654–64.
8. Vallieres M, Zwanenburg A, Badic B, Cheze-Le Rest C, Visvikis D, Hatt M. Responsible Radiomics Research for Faster Clinical Translation. *J Nucl Med*. 2017; jnmed.117.200501.
9. Hatt M, Tixier F, Pierce L, Kinahan PE, Le Rest CC, Visvikis D. Characterization of PET/CT images using texture analysis: the past, the present... any future? *Eur J Nucl Med Mol Imaging*. 2017;44(1):151–65.
10. Orlhac F, Soussan M, Maisonneuve J-A, Garcia CA, Vanderlinden B, Buvat I. Tumor Texture Analysis in 18F-FDG PET: Relationships Between Texture Parameters, Histogram Indices, Standardized Uptake Values, Metabolic Volumes, and Total Lesion Glycolysis. *J Nucl Med*. 2014;55(3):414–22.
11. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges J-P, et al. Intratumor Heterogeneity Characterized by Textural Features on Baseline 18F-FDG PET Images Predicts Response to Concomitant Radiochemotherapy in Esophageal Cancer. *J Nucl Med*. 2011;52(3):369–78.
12. Zwanenburg A. Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. *Eur J Nucl Med Mol Imaging*. 2019;
13. Pfähler E, Jong JR De, Boellaard R. Repeatability of 18 F-FDG PET radiomic features : A phantom study to explore sensitivity to image reconstruction settings , noise , and delineation method. 2018;0(0):1–14.
14. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and Reproducibility of Radiomic Features : A Systematic Review. *Radiat Oncol Biol*. 2018;
15. Galavis P, Hollensen C, Jallow N, Paliwal B, Jeraj R. NIH Public Access. *Acta Oncol*. 2010;49(7):1012–6.
16. Nyflot MJ, Yang F, Byrd D, Bowen SR, Sandison GA, Kinahan PE. Quantitative radiomics: impact of stochastic effects on textural feature analysis implies the need for standards. *J Med Imaging*. 2015;2(4):041002.
17. Boellaard R, Delgado-Bolton R, Oyen WJG, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging [Internet]*. 2015 Feb 2 [cited 2016 Nov 24];42(2):328–54. Available from: <http://link.springer.com/10.1007/s00259-014-2961-x>
18. Machado MAD, Menezes VO, Namias M, Vieira NS, Queiroz CC, Matheoud R, et al. Protocols for Harmonized Quantification and Noise Reduction in Low Dose Oncological ¹⁸F-FDG PET/CT Imaging. *J Nucl Med Technol*. 2018; jnmt.118.213405.
19. Kanoun S, Tal I, Berriolo-riedinger A, Rossi C. Influence of Software Tool and Methodological Aspects of Total Metabolic Tumor Volume Calculation on Baseline [¹⁸F] FDG PET to Predict Survival in Hodgkin Lymphoma. 2015;C:1–15.
20. Mettler J. Metabolic Tumour Volume for Response Prediction in Advanced-Stage Hodgkin Lymphoma. *J Med Imaging*. 2018;1:1–25.
21. Griethuysen J, Fedorov A, Parmar C, Hosny A, Narayan V. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res*. 2018;77(21):1–8.
22. Hatt M, Majdoub M, Vallieres M, Tixier F, Le Rest CC, Groheux D, et al. 18F-FDG PET Uptake Characterization Through Texture Analysis: Investigating the Complementary Nature of Heterogeneity and Functional Tumor Volume in a Multi-Cancer Site Patient Cohort. *J Nucl Med*. 2015;56(1):38–44.
23. Lartizien C, Rogez M, Niaf E, Ricard F. Computer-aided staging of lymphoma patients with FDG PET/CT imaging based on textural information. *IEEE J Biomed Heal Informatics*. 2014;18(3):946–55.
24. Kaalep A, Burggraaff CN, Pieplensbosch S, Verwer EE, Sera T, Zijlstra J, et al. Quantitative implications of the updated EARL 2019 PET – CT performance standards. 2019;1–16.
25. Menezes VO, Machado MAD, Queiroz CC, Souza SO, D'Errico F, Namias M, et al. Optimization of oncological 18 F-FDG PET/CT imaging based on a multiparameter analysis. *Med Phys*. 2016;43(2):930–8.
26. Kaalep A, Sera T, Rijnsdorp S, Yaqub M, Talsma A, Lodge MA, et al. Feasibility of state of the art PET / CT systems performance harmonisation. 2018;1344–61.
27. Ypsilantis P, Siddique M, Sohn H, Davies A. Predicting Response to Neoadjuvant Chemotherapy with PET Imaging Using Convolutional Neural Networks. *PLoS One*. 2015;1–18.
28. Bernardi E De, Buda A, Guerra L, Vicini D, Elisei F, Landoni C, et al. Radiomics of the primary tumour as a tool to improve 18 F-FDG-PET sensitivity in detecting nodal metastases in endometrial cancer. 2018;1–9.
29. Collarino A, Ph D, Garganese G, Fragomeni SM, Arias-LMP, Ph D, et al. Radiomics in vulvar cancer : first clinical experience using 18 F-FDG PET / CT images. 2018;31(0).
30. Zwanenburg, Alex et al. The image biomarker standardisation initiative.
31. Aide N, Lasnon C, Veit-haibach P, Sera T, Sattler B, Boellaard R. EANM / EARL harmonization strategies in PET quantification : from daily practice to multicentre oncological studies. 2017;
32. Orlhac F, Boughdad S, Philippe C, Stalla-Bourdillon H, Nioche C, Champion L, et al. A post-reconstruction harmonization method for multicenter radiomic studies in PET. *J Nucl Med*. 2018;2(Md);jnmed.117.199935.

Contact:

Corresponding author:

Marcos Antônio Dórea Machado

São Rafael Hospital, Nuclear Medicine Department.

Av. São Rafael, 2152 - São Marcos, Salvador/ BA/

Brazil, Zip code: 41253-190

E-mail: machado@radtec.com.br

Artigo nº 3

**Seleção e Harmonização dos Atributos Radiômicos em Tomografia
Computadorizada de COVID-19**

Revista Brasileira de Física Médica

Publicado

Artigo Original

Revista Brasileira de Física Médica (2022) 16:670

Seleção e Harmonização dos Atributos Radiômicos em Tomografia Computadorizada de COVID-19

Radiomic Feature's Selection and Harmonization of COVID-19 in Computed Tomography

Danillo M. Oliveira^{1 2 3}, Marcos A. D. Machado^{4 5}, Mércia L. Oliveira², Ronnyldo R.E. Silva^{5 6}, Mauro Namias⁷, Thamiris R. Reina⁸, Paulo C. O. A. Junior⁸, Andrea S. Lessa⁸, Arquimedes A. B. Lira⁹, Krsna M. A. Rodrigues^{2 3}, Thauan F. Moraes², Vinicius O. Menezes^{5 10}

¹Hospital Universitário de Lagarto – HUL/UFS-EBSERH, Lagarto (SE), Brasil

²Centro Regional de Ciências Nucleares do Nordeste – CRCN-NE, Recife (PE), Brasil

³Departamento de Energia Nuclear – DEN-UFPE, Recife (PE), Brasil

⁴Complexo Hospitalar Universitário Prof. Edgard Santos – UFBA-EBSERH, Salvador (BA), Brasil

⁵Radtec Serviços em Física Médica, Salvador (BA), Brasil

⁶Universidade Federal de Campina Grande - UFCG, Campina Grande (PB), Brasil

⁷Fundación Centro Diagnóstico Nuclear, Buenos Aires (BA), Argentina

⁸Hospital Universitário de Juiz de Fora – HJUF/UFJF-EBSERH, Juiz de Fora (MG), Brasil

⁹Hospital Universitário Alcides Carneiro – HUAC/UFCEG-EBSERH, Campina Grande (PB), Brasil

¹⁰Hospital das Clínicas da Universidade Federal de Pernambuco – HC/UFPE-EBSERH, Recife (PE), Brasil

Resumo

O advento da radiômica tem gerado oportunidades em várias aplicações diagnósticas, incluindo a tomografia computadorizada (TC) de tórax para diagnóstico e estadiamento da COVID-19. Este trabalho demonstra a aplicação de técnicas de seleção e harmonização de atributos radiômicos em um estudo multicêntrico de COVID-19. 37 indivíduos com COVID-19 e 36 indivíduos com outras pneumonias não associadas a COVID-19 foram selecionados de três centros diagnósticos. Noventa e quatro atributos radiômicos da TC de tórax foram quantificados e selecionados usando dois métodos: 1) supervisionado, utilizando análise univariada pela curva ROC, e 2) não supervisionado, utilizando concordância de observadores intraclasse ($ICC > 0,9$) a partir de réplicas das TC com variações na qualidade de imagem e na segmentação. Adicionalmente, foram analisadas características populacionais que pudessem enviesar a harmonização dos atributos. O método de seleção não supervisionado demonstrou ser eficaz na triagem de atributos harmonizados, não sendo necessário uma harmonização com pós processamento. Os atributos selecionados pelo método supervisionado possuem maior dispersão das distribuições das quantificações entre os equipamentos, fazendo-se necessária a aplicação da harmonização ComBat após avaliação da sensibilidade do viés ecológico.

Palavras-chave: Tomografia Computadorizada; COVID-19; Radiômica; Harmonização.

Abstract

The advent of radiomics has emerged with opportunities in various diagnostic applications, including chest computed tomography (CT) for diagnosis and staging of COVID-19. This work demonstrates radiomic selection and harmonization strategies in a multicenter study CT of COVID-19. 37 patients with COVID-19 and 36 patients with other pneumonias not associated with COVID-19 were selected from three diagnostic centers. Ninety-four radiomic features were quantified and selected using two methods: 1) supervised, using univariate analysis by the ROC curve, and 2) unsupervised, using intraclass observer agreement ($ICC > 0.9$) from CT replicates with variations image quality and segmentation. Additionally, population characteristics that could bias feature's harmonization were analyzed. The unsupervised selection method proved to be effective in the screening of harmonized features, without the need for numerical harmonization. Features selected by the supervised method had larger discrepancy among scanner distributions leading to the need of ComBat harmonization after evaluation of ecological bias sensitivity.

Keywords: Computed Tomography; COVID-19; Radiomics; Harmonization.

1. Introdução

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)¹.

Um método de imagem que é recomendado para o diagnóstico da COVID-19 é o exame de Tomografia Computadorizada (TC) de tórax, em que é possível avaliar o comprometimento da doença e também tem se mostrado útil para o prognóstico. Entretanto, este diagnóstico pode ser confundido com o de outras infecções agudas como influenza, SARS, MERS e H1N1².

A radiômica é um método que tem o potencial de extrair características das imagens que estejam associadas com a doença, abrindo uma nova fronteira para o diagnóstico da COVID-19³.

As diferenças de modelos de equipamentos, protocolos de imagens e parâmetros de quantificação podem influenciar os resultados de estudos baseados em radiômica. Apesar do crescente número de publicações envolvendo este método, a seleção de biomarcadores reprodutíveis⁴ e harmonizados^{5,7} são essenciais para garantir a introdução da radiômica na prática clínica.

Uma maneira de viabilizar a intercomparação destes biomarcadores entre diferentes instituições é a aplicação de técnicas de harmonização. No estudo realizado por Orhac e colaboradores (2018)⁵ foi demonstrado que o método de harmonização ComBat 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*) adquiridos em diferentes centros e sob diferentes condições.

Além disso, Orhac e colaboradores (2019)⁶ demonstraram a eficácia do método de harmonização ComBat em exames de tomografia computadorizada, utilizando simuladores de imagens e exames tomográficos reconstruídos com diferentes protocolos.

Este trabalho demonstra dois métodos de seleção de atributos: 1) não supervisionado, com harmonização intrínseca; e 2) supervisionado, com posterior harmonização ComBat. Uma seleção de atributos radiômicos harmonizados são apresentados e sugeridos para estudos de aplicações diagnósticas.

2. Materiais e Métodos

2.1. Requisitos Éticos

Este trabalho seguiu os princípios éticos da Declaração de Helsinque e foi aprovado no Comitê de Ética em Pesquisa das instituições envolvidas:

- Complexo Hospitalar Universitário Prof. Edgard Santos – HUPES/UFBA (4.231.669) - Hospital A;

- Hospital Universitário Alcides Carneiro – HUAC/UFMG (4.818.271) - Hospital B; e
- Hospital Universitário de Juiz de Fora – HUJF/UFJF (4.926.688) - Hospital C.

A assinatura do Termo de Consentimento Livre e Esclarecido pelos participantes da pesquisa foi dispensada.

2.2. Aquisição das Imagens

Os exames de TC de tórax dos pacientes foram adquiridos nos três hospitais com os equipamentos apresentados na Tabela 1.

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).

Foram selecionados retrospectivamente 37 indivíduos com TC de tórax realizados entre abril de 2020 e abril de 2021, com confirmação da COVID-19 no teste RT-PCR. Além disso, foram selecionados 36 indivíduos com TC de tórax com diagnóstico de pneumonia antes da pandemia de COVID-19, realizados entre janeiro de 2018 e outubro de 2019 (Tabela 2).

Tabela 2 - Características dos pacientes

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).

Os critérios de exclusão foram: imagens TC 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*).

2.3. Segmentação das Regiões de Interesse e Extração dos Atributos Radiômicos

Três médicos radiologistas de cada centro de imagem segmentaram manualmente as regiões de interesse - ROI (do inglês - *region of interest*) relacionadas a pneumonias utilizando o software Lifex, versão 6.30 (www.lifexsoft.org). Estas ROIs foram usadas previamente para treinar uma rede neural convolucional 2D de segmentação automática com um esquema de validação cruzada de cinco vezes⁹, onde cada conjunto com 60 pacientes foi usado para treinamento da rede, e os pacientes restantes foram usados para segmentação e extração dos atributos radiômicos. 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. 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 utilizando o PyRadiomics (GRIETHUYSEN, VAN et al., 2017)⁹. A matriz foi reamostrada para voxels de 1x1x2 mm³ e a discretização dentro de cada VOI (do inglês - *volume of interest*) foi dimensionada para 128 níveis de cinza. Um total de 94 atributos foram extraídos para cada lesão, incluindo 18 atributos de primeira ordem e 76 atributos de textura (material suplementar).

2.4. Avaliação do Viés Populacional

Para aplicação do método de harmonização numérica, faz-se necessário que o perfil dos pacientes seja equivalente entre as três instituições, para garantir que os atributos sejam harmonizados apenas quanto ao efeito do equipamento, minimizando vieses relacionados à população estudada (viés ecológico). Desta forma, as distribuições dos volumes das lesões foram analisadas a fim de identificar vieses populacionais. Os atributos fortemente correlacionados com o volume da lesão foram analisados criticamente, com base na hipótese de que o volume é invariável com relação ao equipamento e assim, portanto, um perfil de distribuição dos volumes divergente entre as instituições deverá representar uma característica diferente da população estudada. A avaliação da correlação do atributo com o volume foi realizada pela correlação de Spearman (ρ), onde $\rho > 0,7$ foi considerado fortemente correlacionado¹⁰.

2.5. Seleção e Harmonização dos Atributos Radiômicos

Foram realizados dois métodos de seleção: um método não supervisionado e um método supervisionado.

2.5.1. Método não Supervisionado de Seleção

Para simular diferentes observadores, foram realizadas erosões de 1 mm e 3 mm nos VOI's, resultando em três conjuntos pareados de lesões segmentadas (VOI original, com 1 mm de erosão e

com 3 mm de erosão); e para simular os efeitos de diferentes protocolos de imagens, foram aplicados um filtro de borrimento e um filtro de melhoramento de bordas nas imagens, resultando em três novos conjuntos (imagem original, com filtro de borrimento e com filtro de melhoramento de bordas).

A reprodutibilidade do atributo foi estimada usando o coeficiente de correlação intraclasses – ICC (do inglês - *intra-class correlation coefficient*) em cada conjunto, selecionando os atributos que obtiveram ICC > 0,90 concomitantemente nos métodos de erosão e qualidade de imagem. Este limite do ICC foi usado como exemplo neste trabalho, mas outros limites menos conservadores podem ser adotados.

2.5.2. Método Supervisionado de Seleção

No método supervisionado, 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). O atributo com o menor *p-value* de cada classe foi selecionado.

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) ao qual a VOI foi delineada.

Além disso, o método supervisionado seleciona atributos baseando-se apenas no poder discriminatório quanto ao status da doença, mas não contempla nenhuma etapa relacionada a variações na qualidade da imagem. Por isto, a inclusão de uma etapa de harmonização numérica é proposta após o processo de seleção.

Deste modo, os atributos considerados fracamente correlacionados com o volume da lesão ($\rho < 0,7$) foram harmonizados utilizando o método ComBat de harmonização¹¹ utilizando a biblioteca NeuroCombat disponível em Python (<https://github.com/jmyoung36/neuroCombat.git>).

3. Resultados

Os atributos radiômicos selecionados pelos métodos supervisionado e não supervisionado são apresentados na Tabela 3.

Tabela 3 - Atributos selecionados

Atributo	Nome	Método de Seleção	ICC: Erosão / Filtro	<i>p-value</i>	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 UNIFORMITY NORMALIZED	Supervisionado	0,151 / 0,480	< 0,001	-0,780
07	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

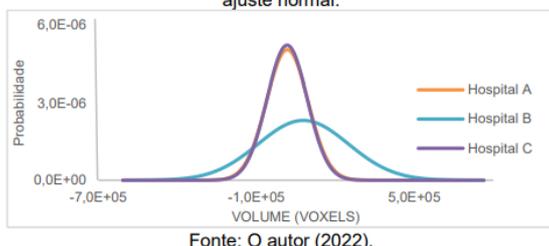
Fonte: O autor (2022).

3.1. Viés Populacional

A análise da Figura 1, que mostra a distribuição dos volumes das lesões (com ajuste para distribuição normal) revela haver uma clara divergência entre os volumes das VOI's dos pacientes do Hospital B em relação aos outros dois hospitais.

A Figura 2 apresenta os gráficos de dispersão entre os valores normalizados (ranking de Besson) dos atributos e dos volumes das lesões. Nota-se que os três atributos selecionados pelo método não supervisionado são fortemente correlacionados com o volume. As figuras 1 e 3 revelam que não há discrepâncias entre as distribuições dos hospitais A e C, enquanto a discrepância observada no hospital B é explicada pela variação populacional deste hospital. Deste modo, não há razão em se aplicar ajustes no domínio dos atributos (harmonização) do método não supervisionado.

Figura 1 - Gráfico da distribuição dos volumes das lesões com ajuste normal.



Fonte: O autor (2022).

Em geral, as correlações com o volume dos atributos extraídos pelo método supervisionado foram fracas (Tabela 3), de modo que o efeito populacional (maior volume no hospital B) pode ser considerado irrelevante na escala destes atributos. Com isto, a harmonização pelo método ComBat ajustará somente o efeito que o equipamento causa na escala do atributo. Dos seis atributos selecionados pelo método supervisionado, apenas o atributo 06 apresentou forte correlação com o volume ($\rho = -0,780$). Entretanto, a figura 2f demonstra que a correlação ocorre com maior força em volumes menores. Note-se que a distribuição deste atributo (figura 4) não apresenta o mesmo padrão das figuras 1 e 3. Desta forma, os seis atributos selecionados pelo método supervisionado foram harmonizados pelo método ComBat.

3.2. Harmonização

3.2.1. Método não Supervisionado

Os atributos selecionados pelo método não supervisionado apresentaram aparente sobreposição entre as distribuições (Figura 3), sugerindo que este método de seleção possui harmonização intrínseca.

3.2.2. Método Supervisionado

Para os atributos selecionados pelo método supervisionado observa-se uma maior divergência entre as distribuições, o que sugere a necessidade de

harmonização numérica. A Figura 4 apresenta os gráficos das distribuições normais dos atributos selecionados pelo método supervisionado sem a harmonização (sem ComBat) e após a harmonização (com ComBat).

4. Discussão

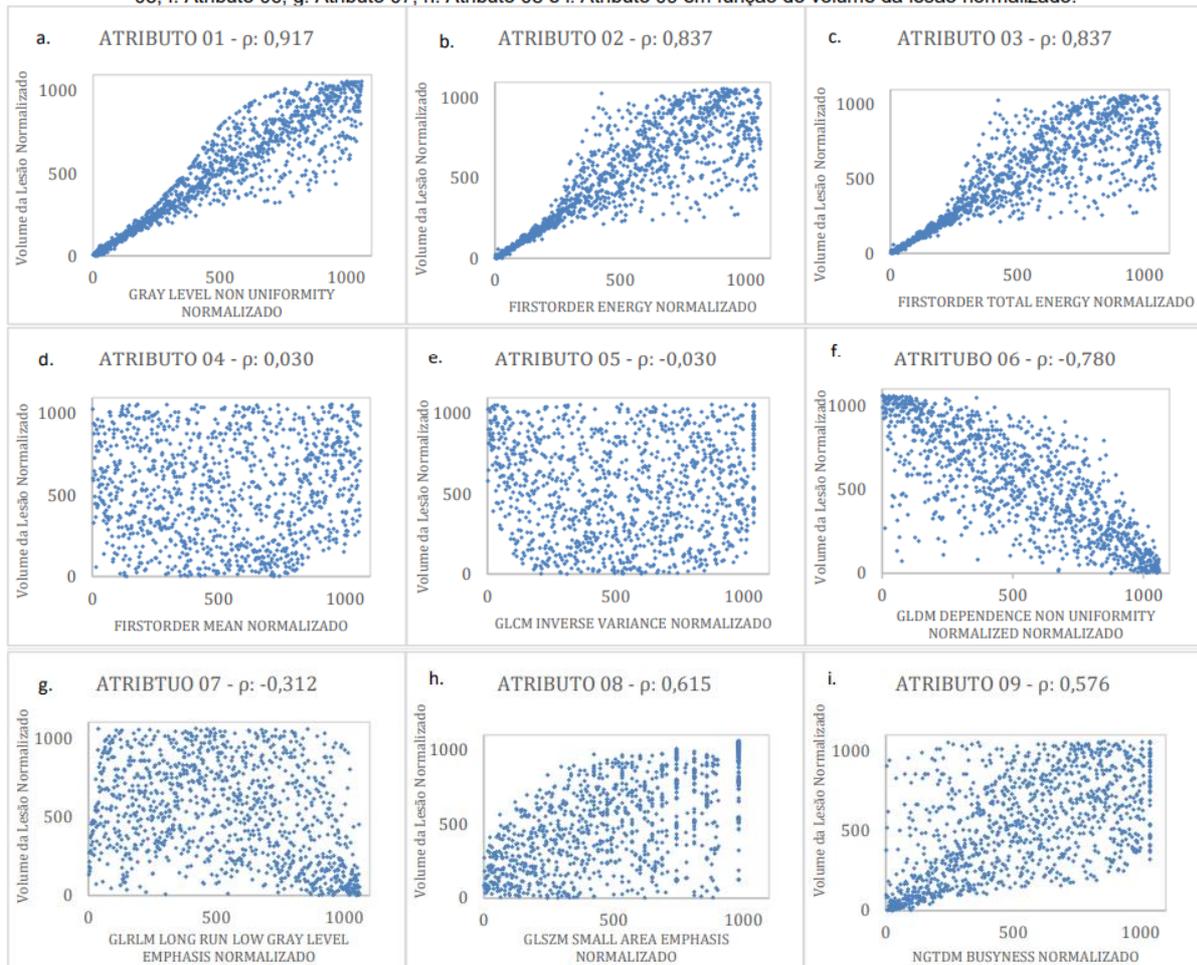
Este trabalho demonstrou uma nova metodologia para seleção e harmonização de atributos radiômicos em TC de tórax de indivíduos infectados com COVID-19. Nove biomarcadores radiômicos foram apresentados como candidatos a aplicações com análises radiômicas.

Orlhac e colaboradores (2019)⁶ utilizaram simuladores de imagens e exames tomográficos de pacientes com câncer de pulmão reconstruídos com diferentes protocolos de TC para avaliar o método ComBat de harmonização. Foram utilizados 10 padrões de textura em um simulador de imagens e 74 pacientes com exames de TC de tórax reconstruídos retrospectivamente com diferentes protocolos, simulando 3 equipamentos. Foi demonstrado que o ComBat removeu o "efeito lote", provocado pelos sistemas de imagens. Este método de harmonização não avaliou a correlação clínica dos atributos com a doença. Entretanto, no presente trabalho a correlação clínica foi avaliada pelo método de seleção supervisionado.

O método de seleção supervisionado avaliou um conjunto de lesões relacionando-as com o status do paciente (COVID-19 / não COVID-19) não sendo possível identificar o VOI que apresentou características da doença, pois um mesmo paciente apresenta lesões com características tanto de COVID-19 como de não COVID-19. Neste caso, pressupõe-se que é mais provável que em um mesmo paciente, a maioria dos VOI's tenham as características da doença apresentada pelo indivíduo. Entretanto, não foi possível diferenciar casos em que lesões tenham sido anotadas erroneamente. Assim, o método supervisionado demonstrou uma associação no nível da lesão, não no nível do paciente. Esta informação encoraja uma ampliação da lista de atributos em novos estudos (material suplementar), ajustando-se os parâmetros ideais pra ICC e *p-value*, visto que valores de ICC menos restritivos podem ser analisados, assim como os marcadores com *p-value* maiores podem se revelar melhores preditores quando analisados no nível do paciente.

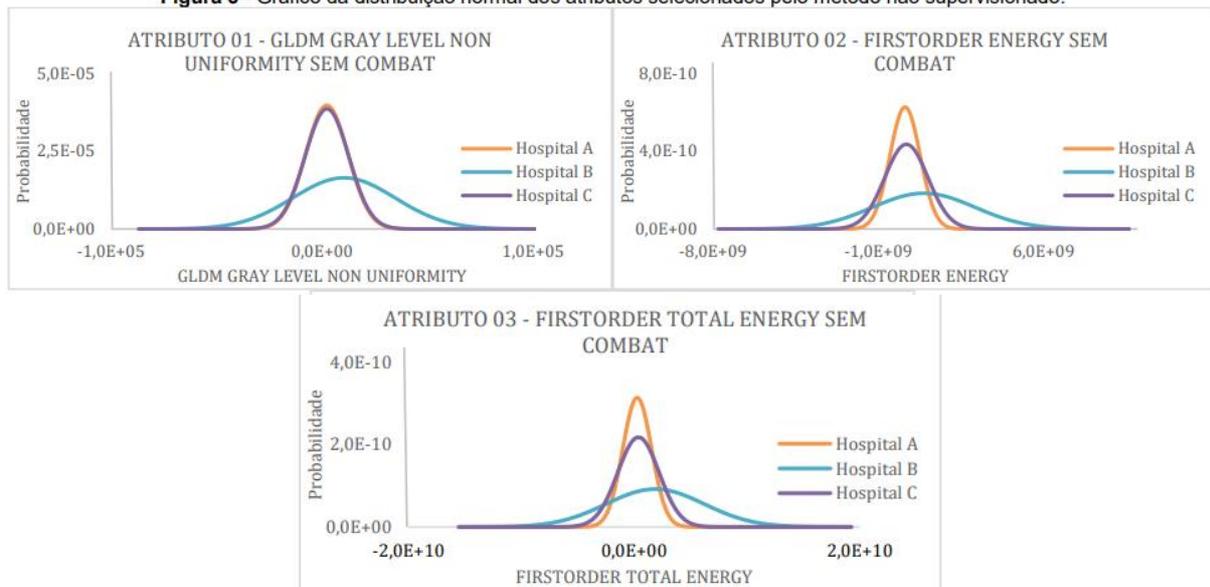
Os atributos selecionados pelo método supervisionado apresentam baixa reprodutibilidade quando expostos a diferenças na segmentação e/ou na qualidade das imagens. Deste modo, é importante que sua aplicação esteja condicionada a sistemas automáticos de segmentação para reduzir o efeito operador-dependente, como também à aplicação de harmonização numérica para equalizar as escalas dos atributos.

Figura 2 - 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.



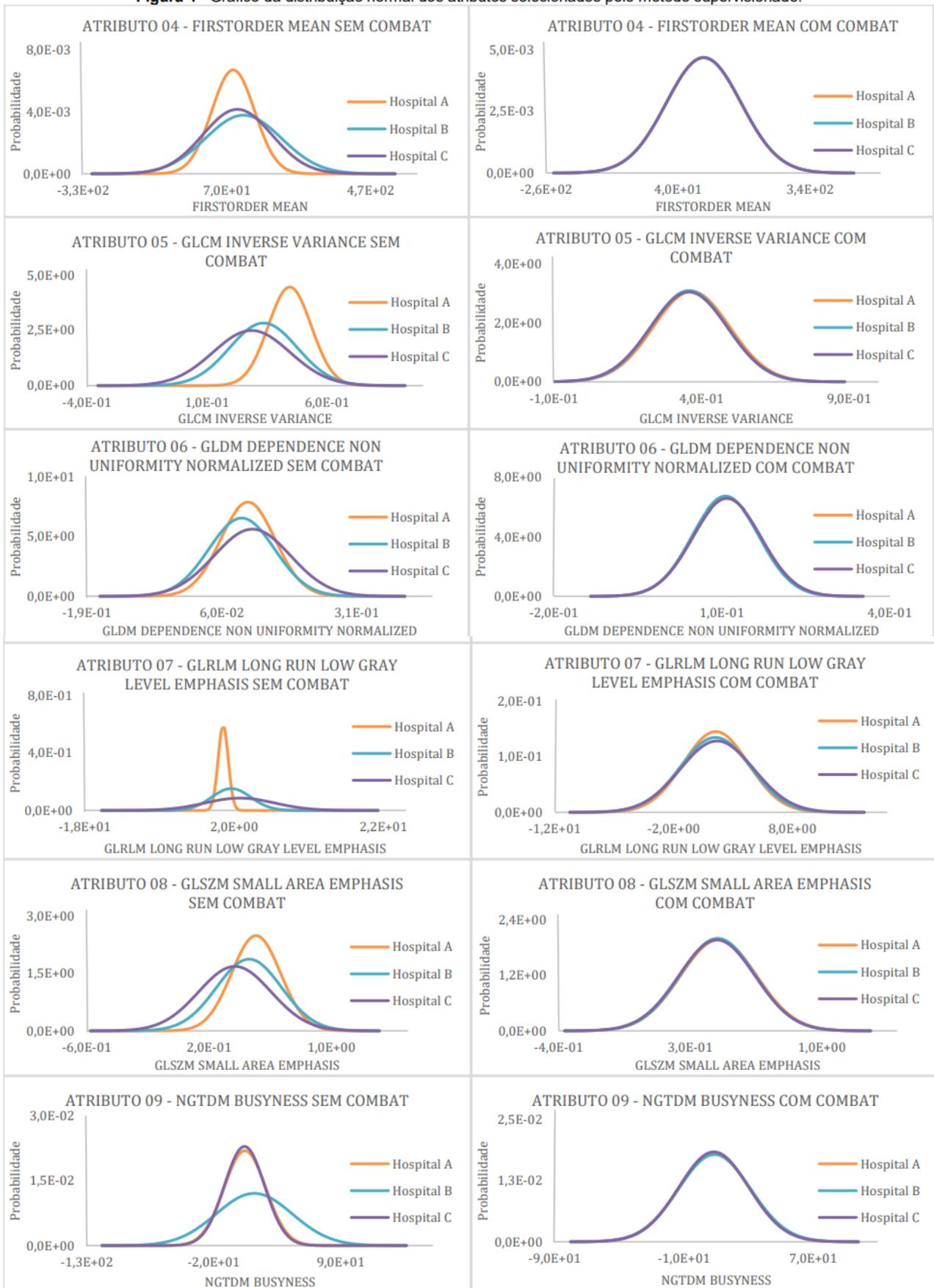
Fonte: O autor (2022).

Figura 3 - Gráfico da distribuição normal dos atributos selecionados pelo método não supervisionado.



Fonte: O autor (2022).

Figura 4 - Gráfico da distribuição normal dos atributos selecionados pelo método supervisionado.



Fonte: O autor (2022).

Fortin e colaboradores (2018)¹² 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.

O presente trabalho possui um viés ecológico 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.

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^{5,7}. 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.

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 borrimento 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.

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.

5. Conclusões

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. A aplicação do método ComBat deve levar em consideração o possível viés ecológico entre as populações estudadas.

Agradecimentos

À Financiadora de Estudos e Pesquisas (Finep) (contrato: 03.20.0156.00); e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior — Brasil (CAPES)— Finance Code 001.

Referências

1. Brasil. Diretrizes Brasileiras para Tratamento Hospitalar do Paciente com COVID-19. Ministério da Saúde. [S.l.]: [s.n.], 2021, p. 1–53.
2. 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.
3. Qingxia Wu, Shuo Wang, Liang Li, Qingxia Wu, Wei Qian, Yahua Hu, Li Li, Xuezhong Zhou, He Ma, Hongjun Li, Meiyun Wang, Xiaoming Qiu, Yunfei Zha and Jie Tian. Radiomics Analysis of Computed Tomography helps predict poor prognostic outcome in COVID-19. *Theranostics* 2020; 10(16): 7231-7244.
4. 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.
5. Orhac, 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.
6. Orhac, 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.
7. Machado, M. A. D.; Menezes, V. O.; Namias, M.; Vieira, N. S.; Queiroz, C. C.; Matheoud, 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.
8. D. P. Fan et al., "Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images," in *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2626-2637, Aug. 2020, doi: 10.1109/TMI.2020.2996645.
9. Van Griethuysen, J. J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Aerts, H. J. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer research*, 77(21), e104-e107.
10. 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>>.
11. 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.
12. 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>>.

Contato:

Nome do Autor Correspondente: *Danillo Menezes Oliveira*

Afiliação: *Hospital Universitário de Lagarto – HUL/UFS – EBSEH*

Endereço: *Av. Brasília, s/n - Novo Horizonte, Lagarto – SE, 49400-000*

E-mail: danillo.oliveira@ebserh.gov.br

Artigo nº 4

**Prediction of Treatment Failure in Hodgkin Lymphoma: A Machine Learning
Radiomic Approach in Baseline ¹⁸F-FDG PET/CT**

Revista Brasileira de Física Médica

Submetido

Prediction of Treatment Failure in Hodgkin Lymphoma: A Machine Learning Radiomic Approach in Baseline 18F-FDG PET/CT

Predição de Falha no Tratamento de Linfoma Hodgkin : Uma Abordagem de Aprendizado de Máquinas em 18F-FDG PET/CT

Marcos A. D. Machado^{1,2} , Cleiton C. Queiroz^{1,3,4}, Jean L. P. Pita⁴, Lucas O. Vieira¹, Bruno C. F. R. Santana, Mauro Namias , Vinicius O. Menezes , Marco A. Salvino, Simone C. S. Brandão, Eduardo M. Netto 

¹Nuclear Medicine Department, São Rafael Hospital, Salvador, Brazil;

²Complexo Hospitalar Universitário Prof. Edgard Santos da Universidade Federal da Bahia/Ebserh, Salvador, Brazil;

³Hospital Universitário Prof. Alberto Antunes da Universidade Federal de Alagoas/Ebserh, Salvador, Brazil;

⁴Department of Computer Science, Universidade Federal de Bahia, Brazil;

⁵Fundación Centro Diagnóstico Nuclear, Buenos Aires, Argentina;

⁶Hospital das Clínicas da Universidade Federal de Pernambuco/Ebserh, Recife, Brazil;

⁷Hematology Department, São Rafael Hospital, Salvador, Brazil

Resumo

Objetivo: A avaliação do estadiamento e prognóstico do Linfoma Hodgkin usando métodos de aprendizado de máquinas ainda é pouco explorado. Este estudo explora características radiômicas e propõe um novo modelo de aprendizado de máquinas para prever falha no tratamento a partir de imagens PET. **Métodos:** Treze características radiômicas foram extraídas de imagens PET de 51 indivíduos com Linfoma Hodgkin com resposta ao tratamento, e 6 indivíduos com falha no tratamento de acordo com o critério de Lugano. Foram feitas análises univariadas das áreas sobre a curva ROC (AUC) para a seleção de características radiômicas, e a correlação de Pearson foi usada para avaliar redundância entre as características radiômicas. Um algoritmo de aprendizado de máquinas foi desenvolvido usando validação cruzada com 6 camadas. O desempenho do algoritmo foi verificado a partir dos resultados das 15 combinações de cada uma das 2 camadas de validação, usando dois métodos (M1 e M2) para classificação do tumor, tanto no nível da lesão quanto no nível do paciente. **Resultados:** As características radiômicas *zone percentage (ZP)*, *high intensity large area emphasis (HILAE)*, *entropy* e *standardized uptake value of maximum pixel (SUVmax)* foram preditores independentes de falha no tratamento. O algoritmo de aprendizado de máquina desempenhou com AUC=0,86 (M1) e AUC=0,96 (M2) para classificar “tumores individuais”, e no “nível paciente” desempenhou com sensibilidade = 80,0%/ especificidade = 88,3% (M1) e sensibilidade = 100%/ especificidade = 100% (M2). O algoritmo desempenhou melhor que o método de estadiamento *Ann Arbor*, que alcançou, respectivamente, 83,3% e 31,7% para sensibilidade e especificidade. **Conclusões:** A avaliação de imagens basais de 18F-FDG PET de pacientes com Linfoma Hodgkin usando aprendizado de máquinas possibilita classificação acurada de pacientes com maior risco de falha no tratamento.

Palavras-chave: imagem molecular, aprendizado de máquinas, linfoma Hodgkin, medicina de precisão.

Abstract

Purpose: Hodgkin Lymphoma staging and prognostic evaluation through radiomic and machine learning (ML) methods has been little explored. This study explores radiomic features and proposes a new ML radiomic model to predict treatment failure in Hodgkin lymphoma patients from PET images. **Methods:** 13 radiomic features were extracted from PET images of 51 subjects with Hodgkin's Lymphoma with response to therapy and 6 subjects with treatment failure as classified by the Lugano criteria for lymphoma response. Univariate analysis was performed for feature selection by means of the areas under ROC curves (AUC), and Pearson correlation was assessed to reduce redundancy in the feature space. A ML algorithm was developed and trained-validated with a six-fold stratified cross validation. The ML performance was assessed through the results of 15 combinations of two-fold validation samples, deriving two ML methods (M1 and M2) for tumor classification at the tumor and patient levels. **Results:** The features *zone percentage (ZP)*, *high intensity large area emphasis (HILAE)*, *entropy* and *standardized uptake value of maximum pixel (SUVmax)* were independent predictors of treatment failure. The ML algorithm performed with AUC=0.86 (M1) and AUC=0.96 (M2) to classify “individual tumors” and at “patient level” performed with sensitivity = 80.0% /specificity = 88.3% (M1) and sensitivity = 100% /specificity = 100% (M2). The ML outperformed the Ann Arbor staging that achieved, respectively, 83.3% and 31.7% for sensitivity and specificity. **Conclusions:** Evaluation of baseline 18F-FDG PET scans of individuals with Hodgkin's lymphoma using a ML radiomic model enabled accurate classification of patients at higher risk of treatment failure.

Keywords: molecular imaging, machine learning, Hodgkin lymphoma, precision medicine.

1. Introduction

Positron emission tomography (PET) and computed tomography (CT) hybrid imaging (PET/CT)

using ¹⁸F-fluorodeoxyglucose (¹⁸F-FDG) holds a major place in staging Hodgkin Lymphoma (HL) with the ability to detect the extent of HL and bone marrow involvement allowing a non-invasive procedure for

staging of nodal and extra-lymphatic disease. Furthermore, ^{18}F -FDG PET/CT also introduced a significant improvement in response evaluation when compared to conventional imaging¹. Additionally to monitoring changes during treatment, recent studies have shown that baseline PET has the potential of predicting response to therapy by measuring the total metabolic tumor volume (TMTV) biomarker^{2,3}.

Classical measurements of tumor uptake (SUV and TMTV) ignore the intratumoral ^{18}F -FDG spatial distribution (i.e.: tumor heterogeneity). The rapidly emerging field of “radiomics” computes quantitative image features that characterize this intratumoral heterogeneity or other tumor phenotypes, leading to a promising field in the era of precision medicine. There is growing evidence that radiomics could add useful information to baseline PET to predict outcome in some types of cancer. Bouallègue *et al* investigated the feasibility of textural and morphological features in predicting early metabolic response assessed from baseline ^{18}F -FDG PET on a heterogeneous cohort of Hodgkin and non-Hodgkin lymphoma patients⁴. To the best of our knowledge, only one study has demonstrated that machine learning (ML) from baseline PET radiomics in HL can predict treatment outcome, achieving nearly 95% accuracy, suggesting that radiomics may enhance the predictive capability of TMTV⁵.

However, this study used a validation sample of 28 subjects and did not allow for final conclusions on ML radiomic prediction of treatment failure since they included only two subjects with treatment failure and early stage HL. On the other hand, Lartzien and colleagues demonstrated the use of radiomics ^{18}F -FDG PET/CT in HL patients to differentiate cancer lesions from hypermetabolic inflammatory foci⁶.

In this work, we explored radiomic features extracted from baseline ^{18}F -FDG PET scans of HL patients and developed a ML classifier to predict treatment failure through a cross-validation methodology.

2. Methods

Scanner, imaging and patient selection

^{18}F -FDG PET/CT imaging was performed on a PET Siemens Biograph TruePoint TrueV (Knoxville, TN, USA) combined with a 16-slice helical CT scanner (Emotion 16; Siemens). PET images were corrected for random coincidences, normalization, dead time losses, scatter and attenuation.

Interim PET (iPET) of HL patients is recommended most often after two cycles to assess early response to therapy. The purpose is to confirm the effectiveness of treatment, using iPET to tailor treatment according to individual response¹⁴. The time of iPET is not mandatory, then clinicians in our institution refer patients for response evaluation during the treatment (iPET) or post therapy (pt-PET), after the treatment is over. Hereafter, we will consider both as pt-PET. Then HL ^{18}F -FDG PET/CT exams that performed both baseline and pt-PET within 12

months from beginning of treatment were selected (mean: 4.5, range: 1.4-11.8). Exclusion criteria were hyperglycemia at the time of tracer administration, or a delay exceeding 90 min between ^{18}F -FDG injection and image acquisition. A total of 57 subjects were selected. An intravenous FDG injection and acquisition duration was performed according to the European Association of Nuclear Medicine (EANM) guidelines⁹ and applying the method previously reported by Menezes *et al.* for noise standardization^{10,11}. Subjects fasted for 6 hours prior to the ^{18}F -FDG injection. The median injection-to-scan time was 66 minutes (standard deviation 10 minutes). Images were reconstructed on a 168x168 matrix size (4.07 x 4.07 mm² voxels) with 3 mm slice thicknesses and using OSEM3D (ordered-subsets expectation maximization) algorithm with 3 iterations, 21 subsets and a 5 mm gaussian filter. These parameters provide quantitatively harmonized images according to EANM.

The study population was selected based on the International Classification of Diseases, Tenth Revision (ICD-10) (C81.X) for HL. The pt-PET was evaluated according to Lugano classification lymphoma response criteria, a score derived from the Deauville 5-point scale (5PS) which scores 1 to 3 are considered Complete Metabolic Response (CMR), score 4 and 5 are considered either No-Metabolic-

Response (NMR), Partial Response (PR), or Progressive Disease (PD) depending on the evolution compared to baseline evaluation (no change from baseline, reduced uptake compared to baseline, and increased uptake or new lesion, respectively). This classification results in standardized reports with good prognosis discrimination and good inter-observer reproducibility¹. We also conducted a review of each patient's electronic medical record to determine their initial status and outcome at pt-PET.

The study was approved by the Hospital São Rafael ethical board (CAAE: 79258017.6.0000.0048) that waived the informed consent, all information was anonymized. The images were acquired according to the clinical protocol of Hospital São Rafael for tumor PET imaging.

Radiomic feature extraction

Individual lesions from baseline PET scans were segmented with the Beth Israel PET/CT plugin for FIJI (ImageJ, Bethesda, MD, USA)^{2,12} using 41% threshold of maximum SUV (SUV_{max}). Segmented VOIs were analyzed with PyRadiomics, an open-

source platform available at www.radiomics.io that enables the extraction of a large panel of radiomic features¹³. We resampled the matrix grid to cubic voxels of 4x4x4 mm³ and used intensity discretization with a fixed bin size of 0.25 g/ml to avoid feature instability¹⁴. The radiomic feature classes and corresponding features are represented in Table 1 as defined by Griethuysen *et al*¹³.

Entropy and difference average from GLCM were calculated according to 2 different methods as

described in Hatt et al. (2015) which will be denoted as E13 and D13, where E is entropy and D is difference average, whereas the parameters from the second method will be denoted as E1 and D1. Thus, a total of 13 features from five classes were analyzed. These features were selected from previous studies^{4-6, 14-17}.

Feature analysis

Lesions with an MTV < 5 cm³ were excluded from the analysis because of insufficient voxels to compute textural indexes. In pt-PET, patients diagnosed as PD represent the group at higher risk of treatment failure and death^{5,7,8}. Therefore, we classified all VOIs within the baseline ¹⁸F-FDG PET/CT as PD for such patient (non-responder group). For patients diagnosed as CMR, NMR and PR in pt-PET we categorized all baseline VOIs as CMR (responder group).

Statistical analysis

Tumor level: Analysis for association between features and response to therapy status were performed by means of the area under the ROC curve (AUC) computation. Features with p value < 0.05 were selected and Pearson correlation was used to reduce the space by selecting features with $r < 0.90$. The population was randomly divided into 6 groups. All tumors from four groups were used to train a random forest based AdaBoost enhanced machine learning (ML) model to discriminate tumors into two classes (progressive disease vs. non-progressive disease) and the two remaining groups of tumors were used for model validation. To estimate the classifier performance, we used a 15-fold cross-validation, using data from two validation groups per fold, each time performing model building with a different training set, and then testing this model on the data from the two unseen validation groups (supplementary material). The Python library Scikit-learn was used to implement these methods, classifying tumors through a radiomic score ranging from 0 to 1, with 0 meaning non-progressive disease and 1 meaning progressive disease¹⁸.

Subject level: Prediction to discriminate responders and non-responders at the subject level was assessed through the tumor with the maximum ML score within each patient, assessing two methods: **M1**) direct computation of sensitivity and specificity at every independent fold: subjects which scored higher than radiomic score threshold S were considered as progressive disease. The sensitivities and specificities were averaged among all folds; and **M2**) averaging the scores from the 15 folds for each tumor (S_m), thus analyzing all 57 subjects together. Subjects which scored higher than S_m were considered as progressive disease. The sensitivity and specificity were computed for the whole dataset.

S and S_m were determined through the Youden's index. The statistical power $1-\beta$ of each method was assessed with significance level $\alpha = 5\%$. Fisher's exact test was performed to verify differences

between groups for categorical variables, and Wilcoxon Mann-Whitney test for continuous data. Also, confidence intervals of sensitivities and specificities were estimated using binomial proportion Wilson's score method. Prognostication was compared with Ann Arbor staging for initial stage (I or II) and advanced stage (III or IV). SPSS 21.0 for Windows or OpenEpi (Emory University) was used for statistical analysis.

3. Results

Population characteristics are presented in Table 2.

Tumor level

Feature selection (Table 3) and reduction (Figure 1), where SUV_{max} , D1 and D13 were highly correlated and were further represented by SUV_{max} alone, revealed five radiomic predictors of treatment failure. ML algorithm thus constructed with these predictors and validated in every 15-independent fold accomplished superior discrimination than individual features. Detailed performance at every 15-validation fold is presented in supplementary material, where the mean AUC of M1 at the lesion level, $AUC_m = 0.86$. For improved readability and interpretation of results, the mean ROC curve for M1 (ROC_m) was simulated by iteratively misspecifying the true values with a random gaussian function. The procedure was stopped when the simulated $AUC_s = AUC_m = 0.86$ ¹⁹. Figure 2 demonstrates performance of univariate features and ML classifiers at the tumor level M1 ($AUC_s = 0.86$, 95% CI: 0.77-0.94) and M2 ($AUC = 0.96$, 95% CI: 0.93-1.00).

Subject level

Upon translating the classification to the subject level, the sensitivity and specificity were, respectively, 80.0% (CI = 43.7-97.0%) and 88.3% (76.6-94.5%) for M1 ($S = 0.55$), and 100% (70.0-100.0%) and 100% (93.0-100%) for M2 ($S_m = 0.58$) (Table 4). Figure 3 illustrates pretreatment radiomic scores among groups of pt-PET status at the tumor level (A) and subject level (B) for M2. Both methods M1 and M2 were strong predictors of treatment failure which the estimated statistical power $1-\beta$ were 94% and 100%, respectively.

Ann Arbor staging had 83.3% (CI = 43.7-97.0%) and 31.7% (CI = 20.3-45.0%) sensitivity and specificity, respectively. Contrary to the ML scores, it was not possible to detect a difference between groups for Ann Arbor staging ($p = 0.66$).

An example of radiomic characteristics of two lesions in Figure 3-B were ML score = 0.77 ($SUV_{max} = 3.7$, ZP = 0.17, Complexity = 11, HILAE = 150 and E1 = 4.5) and ML score = 0.30 ($SUV_{max} = 7.8$, ZP = 0.54, Complexity = 744, HILAE = 4.7 and E1 = 8.4).

4. Discussion

We evaluated newly risk factors of treatment failure and predictive models through baseline ¹⁸F-FDG PET/CT of HL by characterizing tumor ¹⁸F-FDG distribution. We conducted a careful methodological design and data quality control to avoid the possible pitfalls in the radiomic workflow and analysis, and

assure reliability and repeatability of the results^{14,17,19,20}.

Recent studies of morphological biomarker TMTV in baseline ¹⁸F-FDG PET/CT demonstrated that a higher TMTV indicates worst prognosis^{2,3}. Hatt et al (2015) studied associations between textural features and MTV in various tumors, demonstrating the added contribution and dependence of textural features on MTV¹⁵. Here, we evaluated tumor texture as an indicator of disease progression and excluded MTV in our ML model since the study protocol anticipated exclusion of features with $p \geq 0.05$ for the AUC. Current evidences using radiomics with HL ¹⁸F-FDG PET/CT suggests that robust statistical approaches based on ML would potentially improve predictive models for HL staging^{5,20}. Thus, we conducted two ML radiomic classification methods with good performance using a different approach from Milgrom et al by testing all six “progressive disease” patients through the independent validation folds.

Radiomic features and tumors have a complex relationship. The exact relationship between radiomic features and underlying tumor biology can be established only on carefully designed prospective studies. However, we can assume that a combination of features may be associated with tumor cellularity, vascularization, perfusion, proliferation, aggressiveness, hypoxia, angiogenesis, and necrosis, factors related to more aggressive behavior, poorer response to treatment, and worse prognosis^{22,23}. We demonstrated that five features are informative regarding differences in ¹⁸F-FDG uptake heterogeneity in HL ¹⁸F-FDG PET/CT. From table 3, texture features HILAE, ZP, E1 and Complexity suggests that more homogeneous tumors are less likely to respond to therapy than more heterogeneous ones²⁴. In this study, M1 was assessed through every 15 independent folds, while M2 was not validated in an independent dataset. Then, M2 may have positively biased and lacks internal validation¹⁷. Additionally, feature selection and reduction should ideally be performed using only the model development dataset¹⁷, but it was not possible in this study due to our cross-validation approach using all six “progressive disease” patients. Indeed, results of M1 and M2 are encouraging to validate or refute this pilot study in an external cohort.

Our study has other limitations. First, we used the Lugano Criteria as the gold standard to define treatment failure which has 73% sensitivity and 94% specificity for iPET^{7,8}. Although one would expect pt-PET is more accurate and meaningful than iPET for prognostication, selection bias might have occurred in the tumor level, since some tumors in the non-responder group could not present the same radiomic characteristics as others in a more advanced stage of the disease. However, the high specificity of our model could diminish this impact while translating the diagnostic information to the subject level. Second, the study involved only one center and PET system, lacking external validation. However, differences in scanner, reconstruction, and segmentation may be

taken into account by considering the most robust features ZP, HILAE, SUV_{max}, E1 and Complexity which have shown robustness in test–retest reproducibility^{15,16}. It is important to highlight that textural features are dependent on the segmentation method (here 41% threshold of SUV_{max}, which increases the weight of regions with high uptake) and the results might be probably different with another segmentation method. Last, we used a small sample size, despite it was sufficiently large to detect differences in radiomic profile among subject groups and powerful enough to model the complex relationship among features. Furthermore, the proposed method outperformed the current Ann Arbor staging.

Our efforts are in line with a global demand to make radiomics and molecular imaging feasible and translational to clinical practice; as methodological procedures, image noise standardization and feature harmonization is becoming more popular^{25,26}.

5. Conclusions

Radiomic features from baseline ¹⁸F-FDG PET scans of HL are associated with treatment failure by Lugano Criteria and with poorer prognosis. This association enabled a radiomic ML classification that offer promise for personalized medicine, thus stratifying patients to conduct the best treatment strategies.

Acknowledgements

We thank Salim Kanoun for helpful instructions on managing the PET/CT Viewer Beth Israel plugin for FIJI and integrations with Pyradiomics.

Funding: This work was supported in part by the Federal University of Bahia (Permanecer project and Programa de Pós-Graduação em Medicina e Saúde, Faculdade de Medicina da Bahia, Salvador, Bahia, Brazil) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES).

References

1. Kanoun S, Rossi C, Casasnovas O. [¹⁸F]FDG-PET/CT in hodgkin lymphoma: Current usefulness and perspectives. *Cancers (Basel)*. 2018;10(5):1-11. doi:10.3390/cancers10050145
2. Mettler J. Metabolic Tumour Volume for Response Prediction in Advanced-Stage Hodgkin Lymphoma. *J Med Imaging*. 2018;1:1-25. doi:10.2967/jnumed.118.210047
3. Cottreau AS, Versari A, Loft A, et al. Prognostic value of baseline metabolic tumor volume in early-stage Hodgkin lymphoma in the standard arm of the H10 trial. *Blood*. 2018;131(13):1456-1463. doi:10.1182/blood-2017-07-795476
4. Bouallegue FB, Tabaa YA, Kafrouni M, Cartron G, Vauchot F, Mariano-Goulart D. Association between textural and morphological tumor indices on baseline PET-CT and early metabolic response on interim PET-CT in bulky malignant lymphomas. *Med Phys*. 2017;44(9):4608-4619. doi:10.1002/mp.12349
5. Milgrom SA, Elhalawani H, Lee J, et al. A PET Radiomics Model to Predict Refractory Mediastinal Hodgkin

- Lymphoma. 2019;(December 2018):1-8.
doi:10.1038/s41598-018-37197-z
6. Lartzien C, Rogez M, Niaf E, Ricard F. Computer-aided staging of lymphoma patients with FDG PET/CT imaging based on textural information. *IEEE J Biomed Heal Informatics*. 2014;18(3):946-955.
doi:10.1109/JBHI.2013.2283658
 7. Barrington SF, Mikhaeel NG, Kostakoglu L, et al. Role of imaging in the staging and response assessment of lymphoma: Consensus of the international conference on malignant lymphomas imaging working group. *J Clin Oncol*. 2014;32(27):3048-3058.
doi:10.1200/JCO.2013.53.5229
 8. Biggi A, Gallamini A, Chauvie S, et al. International Validation Study for Interim PET in ABVD-Treated, Advanced-Stage Hodgkin Lymphoma: Interpretation Criteria and Concordance Rate Among Reviewers. *J Nucl Med*. 2013;54(5):683-690.
doi:10.2967/jnumed.112.110890
 9. Boellaard R, Delgado-bolton R, Oyen WJG, et al. FDG PET / CT : EANM procedure guidelines for tumour imaging : version 2 . 0. 2015:328-354.
doi:10.1007/s00259-014-2961-x
 10. Menezes V, Machado M, Queiroz C, et al. How does Lean Six Sigma method improve healthcare practice in nuclear medicine departments? A successful case of dedicated software applications in oncological PET/CT. *J Nucl Med*. 2018;59(supplement 1):1013-1013.
http://jnm.snmjournals.org/cgi/content/short/59/supplement_1/1013. Accessed April 16, 2019.
 11. Menezes VO, Machado MAD, Queiroz CC, et al. Optimization of oncological 18 F-FDG PET/CT imaging based on a multiparameter analysis. *Med Phys*. 2016;43(2):930-938. doi:10.1118/1.4940354
 12. Kanoun S, Tal I, Berriolo-riedinger A, Rossi C. Influence of Software Tool and Methodological Aspects of Total Metabolic Tumor Volume Calculation on Baseline [18F] FDG PET to Predict Survival in Hodgkin Lymphoma. 2015;C:1-15. doi:10.1371/journal.pone.0140830
 13. Griethuysen J, Fedorov A, Parmar C, Hosny A, Narayan V. HHS Public Access. *Cancer Res*. 2018;77(21):1-8. doi:10.1158/0008-5472.CAN-17-0339.Computational
 14. Pfaehler E, Jong JR De, Boellaard R. Repeatability of 18 F-FDG PET radiomic features : A phantom study to explore sensitivity to image reconstruction settings , noise , and delineation method. 2018;0(0):1-14.
doi:10.1002/mp.13322
 15. Hatt M, Majdoub M, Vallieres M, et al. 18F-FDG PET Uptake Characterization Through Texture Analysis: Investigating the Complementary Nature of Heterogeneity and Functional Tumor Volume in a Multi-Cancer Site Patient Cohort. *J Nucl Med*. 2015;56(1):38-44. doi:10.2967/jnumed.114.144055
 16. Galavis P, Hollensen C, Jallow N, Paliwal B, Jeraj R. NIH Public Access. *Acta Oncol*. 2010;49(7):1012-1016.
doi:10.3109/0284186X.2010.498437
 17. Zwanenburg A. Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. *Eur J Nucl Med Mol Imaging*. 2019. doi:10.1007/s00259-019-04391-8
 18. Pedregosa F, Weiss R, Brucher M. Scikit-learn : Machine Learning in Python. 2011;12:2825-2830.
 19. Kundu S, Kers JG, Janssens ACJW. Constructing Hypothetical Risk Data from the Area under the ROC Curve : Modelling Distributions of Polygenic Risk. 2016:1-11. doi:10.1371/journal.pone.0152359
 20. Hatt M, Tixier F, Pierce L, Kinahan PE, Le Rest CC, Visvikis D. Characterization of PET/CT images using texture analysis: the past, the present... any future? *Eur J Nucl Med Mol Imaging*. 2017;44(1):151-165.
doi:10.1007/s00259-016-3427-0
 21. Vallieres M, Zwanenburg A, Badic B, Cheze-Le Rest C, Visvikis D, Hatt M. Responsible Radiomics Research for Faster Clinical Translation. *J Nucl Med*. 2017;jnumed.117.200501.
doi:10.2967/jnumed.117.200501
 22. Cook GJR, Yip C, Siddique M, et al. Are Pretreatment 18F-FDG PET Tumor Textural Features in Non-Small Cell Lung Cancer Associated with Response and Survival After Chemoradiotherapy? *J Nucl Med*. 2013;54(1):19-26. doi:10.2967/jnumed.112.107375
 23. Tixier F, Le Rest CC, Hatt M, et al. Intratumor Heterogeneity Characterized by Textural Features on Baseline 18F-FDG PET Images Predicts Response to Concomitant Radiochemotherapy in Esophageal Cancer. *J Nucl Med*. 2011;52(3):369-378.
doi:10.2967/jnumed.110.082404
 24. Coroller TP, Agrawal V, Huynh E, et al. Radiomic-Based Pathological Response Prediction from Primary Tumors and Lymph Nodes in NSCLC. *J Thorac Oncol*. 2017;12(3):467-476. doi:10.1016/j.jtho.2016.11.2226
 25. Orihac F, Boughdad S, Philippe C, et al. A post-reconstruction harmonization method for multicenter radiomic studies in PET. *J Nucl Med*. 2018;2(Md);jnumed.117.199935.
doi:10.2967/jnumed.117.199935
 26. Sollini M, Cozzi L, Antunovic L, Chiti A, Kirienko M. PET Radiomics in NSCLC: State of the art and a proposal for harmonization of methodology. *Sci Rep*. 2017;7(1):1-15.
doi:10.1038/s41598-017-00426-y

Contact:

Marcos Machado
Nuclear Medicine Department,
Hospital São Rafael, Avenida São Rafael, 2152
São Marcos, Salvador - BA, 41253-190, Brazil.
Tel. +55 71 3281-6891. Fax. +55 71 3281-6327.
E-mail: machado@radtec.com.br

Table 1: Radiomic features.

Class	Features
First-order statistics	SUV _{max} , kurtosis
Shape descriptors	Metabolic tumor volume (MTV), sphericity, surface volume ratio (SVR)
Gray Level Co-occurrence Matrix (GLCM)	Entropy (E), difference average (D)
Gray Level Size Zone Matrix (GLSZM)	Zone percentage (ZP), high intensity large area emphasis (HILAE)
Neighboring Gray Tone Difference Matrix (NGTDM)	Coarseness, complexity

Table 2: Population characteristics

Characteristics	Responders (n=51)	Non-responders (n=6)	p-value
Median age (range)	28 (4-82)	18 (7-29)	0.04
Male	53%	50%	1.00
Ann Arbor			0.66
stage I/ stage II	2% /29%	0% /17%	
stage III/ stage IV	26% /43%	33% /50%	
B symptoms	88% ^a	67% ^b	0.40
Bulk (>10 cm)	59%	50%	0.69
Extra nodal disease	65%	50%	1.00
ABVD or ABVD-like chemotherapy	100%	100%	1.00
Median time (month) §	4.7 (1.4-11.3)	4.5 (1.8-11.8)	0.94

Valid values computed: 10 subjects with missing values^a, 3 subjects with missing values^b.

§ Time period between baseline and pt-PET

Table 3: Feature statistics in the lesion level

Features	Responder (n=289)		Non-responder (n=31)		p-value
	IQR	Median	IQR	Median	
MTV (cm ³)	8.8-31.7	14.7	7.6-38.7	15.7	ns
SUV _{max}	6.6-12.1	8.9	4.6-8.5	7.1	**
Kurtosis	2.4-3.5	2.8	2.5-3.2	2.8	ns
D13	3.1-6.0	4.5	1.8-4.0	3.4	**
D1	3.0-6.2	4.4	1.7-3.9	3.4	**
E13	6.3-7.8	7.0	5.5-7.8	6.4	ns
E1	7.1-8.9	7.9	5.7-8.0	7.6	***
ZP	0.36-0.63	0.53	0.24-0.47	0.41	**
HILAE	3.8-19.0	6.7	10.5-116.3	17.0	**
Coarseness	0.01-0.04	0.03	0.01-0.05	0.03	ns
Complexity	286-1844	720	52-611	349	**
Sphericity	0.53-0.70	0.63	0.54-0.75	0.65	ns
SVR	0.26-0.38	0.32	0.25-0.39	0.32	ns

IQR: interquartile range; MTV: metabolic tumor volume; SUV_{max}: standardized uptake value of maximum pixel; D13: difference average using 13 matrices; D1: difference average using 1 matrix; E13: entropy using 13 matrices; E1: entropy using one matrix; ZP: zone percentage; HILAE: high intensity large area emphasis; SVR: surface volume ratio; p-value of AUC: ** p ≤ 0.001, *** 0.001 < p < 0.05, ns: p ≥ 0.05.

Figure 1: Correlation matrix of best features predictors.

D13	0.94					
HILAE	-0.07	-0.06				
ZP	0.59	0.67	-0.13			
Complexity	0.81	0.85	-0.01	0.23		
D1	0.95	0.98	-0.06	0.67	0.82	
E1	0.56	0.56	-0.12	0.75	0.18	0.50
	SUVmax	D13	HILAE	ZP	Complexity	D1

Features with Pearson correlation $r > 0.9$ are highlighted.

Figure 2: ROC curves of univariate features and machine learning classifier.

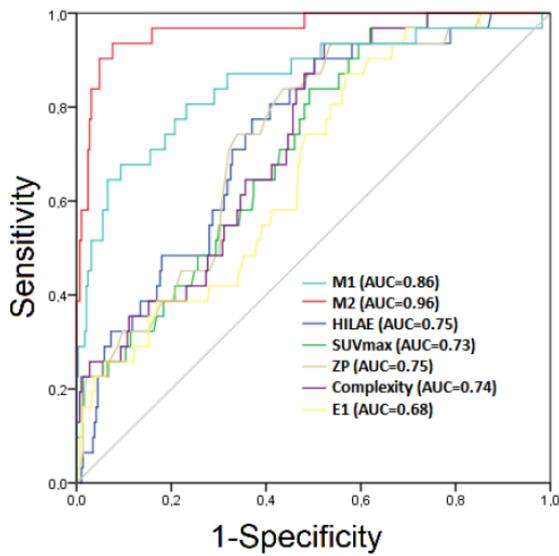
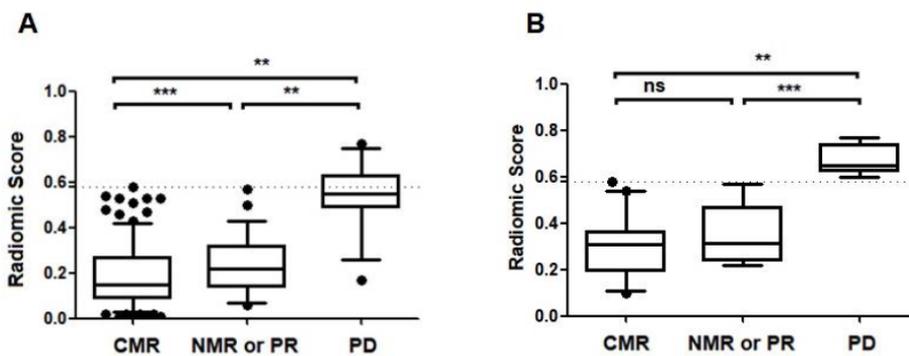


Figure 3: Radiomic score of M2 versus pt-PET status.



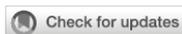
A: scores for individual tumors. B: maximum tumor score for each subject. Boxplots represents 5-95% distribution. Dashed lines represent optimal score threshold of 0.58. CMR: complete metabolic response, NMR or PR: no metabolic response or partial response, and PD: progressive disease. ** $p < 0.001$. *** $p < 0.05$. ns: $p \geq 0.05$.

Artigo n° 5

Exploratory analysis of radiomic as prognostic biomarkers in 18F-FDG PET/CT scan in cervical cancer

Frontiers in Medicine

Publicado



OPEN ACCESS

EDITED BY
 Giorgio Treglia,
 Ente Ospedaliero Cantonale (EOC),
 Switzerland

REVIEWED BY
 Paola Mapelli,
 Vita-Salute San Raffaele University, Italy
 Laura Evangelista,
 University of Padua, Italy

*CORRESPONDENCE
 Nadja Rolim Gonçalves de Alencar
 nadja.rolim@gmail.com

SPECIALTY SECTION
 This article was submitted to
 Nuclear Medicine,
 a section of the journal
 Frontiers in Medicine

RECEIVED 16 September 2022
 ACCEPTED 10 November 2022
 PUBLISHED 02 December 2022

CITATION
 Alencar NRG, Machado MAD,
 Mourato FA, Oliveira ML, Moraes TF,
 Mattos Junior LAR, Chang TC,
 Azevedo CRAS and Brandão SCS
 (2022) Exploratory analysis of
 radiomic as prognostic biomarkers in
¹⁸F-FDG PET/CT scan in uterine
 cervical cancer.
Front. Med. 9:1046551.
 doi: 10.3389/fmed.2022.1046551

COPYRIGHT
 © 2022 Alencar, Machado, Mourato,
 Oliveira, Moraes, Mattos Junior,
 Chang, Azevedo and Brandão. This is
 an open-access article distributed
 under the terms of the Creative
 Commons Attribution License (CC BY).
 The use, distribution or reproduction in
 other forums is permitted, provided
 the original author(s) and the copyright
 owner(s) are credited and that the
 original publication in this journal is
 cited, in accordance with accepted
 academic practice. No use, distribution
 or reproduction is permitted which
 does not comply with these terms.

Exploratory analysis of radiomic as prognostic biomarkers in ¹⁸F-FDG PET/CT scan in uterine cervical cancer

Nadja Rolim Gonçalves de Alencar^{1,2*},
 Marcos Antônio Dórea Machado³, Felipe Alves Mourato^{1,2},
 Mércia Liane de Oliveira⁴, Thauan Fernandes Moraes⁴,
 Luiz Alberto Reis Mattos Junior⁵, Tien-Man Cabral Chang⁶,
 Carla Rameri Alexandre Silva de Azevedo⁷ and
 Simone Cristina Soares Brandão^{1,2,5}

¹Master of Science Surgery Post-Graduation Program, Federal University of Pernambuco, Recife, Pernambuco, Brazil, ²Department of Radiology and Nuclear Medicine, Hospital das Clínicas, Federal University of Pernambuco, Recife, Pernambuco, Brazil, ³Department of Radiology, Complexo Hospitalar Universitário Professor Edgard Santos/Universidade Federal da Bahia (UFBA), Salvador, Bahia, Brazil, ⁴Northeast Center for Strategic Technologies, Recife, Pernambuco, Brazil, ⁵Clinical Medicine, Center for Medical Sciences, Federal University of Pernambuco, Recife, Pernambuco, Brazil, ⁶Nuclear Medicine Service, Instituto de Medicina Integrada Fernandes Figueira, Recife, Pernambuco, Brazil, ⁷Department of Oncology, Instituto de Medicina Integrada Fernandes Figueira, Recife, Pernambuco, Brazil

Objective: To evaluate the performance of ¹⁸F-fluorodeoxyglucose positron emission tomography (¹⁸F-FDG PET/CT) radiomic features to predict overall survival (OS) in patients with locally advanced uterine cervical carcinoma.

Methods: Longitudinal and retrospective study that evaluated 50 patients with cervical epidermoid carcinoma (clinical stage IB2 to IVA according to FIGO). Segmentation of the ¹⁸F-FDG PET/CT tumors was performed using the LIFEx software, generating the radiomic features. We used the Mann–Whitney test to select radiomic features associated with the clinical outcome (death), excluding the features highly correlated with each other with Spearman correlation. Subsequently, ROC curves and a Kaplan–Meier analysis were performed. A *p*-value < 0.05 were considered significant.

Results: The median follow-up was 23.5 months and longer than 24 months in all surviving patients. Independent predictors for OS were found—SUVpeak with an AUC of 0.74, sensitivity of 77.8%, and specificity of 72.7% (*p* = 0.006); and the textural feature gray-level run-length matrix GLRLM_LRLGE, with AUC of 0.74, sensitivity of 72.2%, and specificity of 81.8% (*p* = 0.005). When we used the derived cut-off points from these ROC curves (12.76 for SUVpeak and 0.001 for GLRLM_LRLGE) in a Kaplan–Meier analysis, we can see two different groups (one with an overall survival probability of approximately 90% and the other with 30%). These biomarkers are independent of FIGO staging.

Conclusion: By radiomic ^{18}F -FDG PET/CT data analysis, SUVpeak and GLRLM_LRLGE textural feature presented the best performance to predict OS in patients with cervical cancer undergoing chemo-radiotherapy and brachytherapy.

KEYWORDS

positron emission tomography, prognosis, uterine cervical neoplasms, ^{18}F -fluorodeoxyglucose, radiomics

Introduction

Cervical uterine cancer is an important cause of death in women, especially in regions of low socioeconomic development (1–3). In more advanced stages, fluorine-18-labeled fluorodeoxyglucose positron emission tomography associated with computed tomography (^{18}F -FDG PET/CT) is recommended for the adequate evaluation of lymph nodes and distant metastases (4–6).

The standardized uptake value (SUV) of ^{18}F -FDG is the most used semi-quantitative variable in ^{18}F -FDG PET/CT (7). This value translates the lesion glycolytic metabolism and the higher the value, the more aggressive the tumor (8). Other quantitative metrics extracted from the ^{18}F -FDG PET/CT scan are the metabolic tumor volume (MTV), which translates the measure of the tumor volume with a higher metabolism, and the total lesion glycolysis rate (TLG), which is the product of the mean SUV by the lesion MTV (9). These three variables reflect the tumor metabolic load and could help to predict the patient's prognosis (10, 11).

Radiomic is the extraction of mineable data from medical imaging that has emerged recently (12). It analyzes the lesion phenotype using n quantifying and characterizing several tumoral features (13–15). Among the numerous variables of the radiomic analysis of ^{18}F -FDG PET/CT images, the textural features present a greater correlation with the heterogeneous biological behavior of the tumor. They may serve as predictive markers of overall survival (OS) and therapeutic response (16–18).

Therefore, this paper aims to find radiomic features and metabolic parameters predictive of OS from ^{18}F -FDG PET/CT scans of uterine cervical cancer.

Materials and methods

Patients and methods

The present study included 50 consecutive patients with histologically confirmed diagnoses of uterine cervical squamous cell carcinoma between 2013 and 2015 (Table 1).

The inclusion criteria were: women over 18 years that were undergone pretreatment ^{18}F -FDG PET/CT. All patients received standardized chemotherapy treatment with cisplatin and gemcitabine, with two cycles of neoadjuvant chemotherapy, with subsequent radiotherapy and brachytherapy according to the institutional protocol. The patients were followed up for at least 24 months. The exclusion criteria included ^{18}F -FDG PET/CT scans in disagreement with the acquisition, processing, or reconstruction parameters, according to the Image Biomarker Standardization Initiative (IBSI) (19). The selected patients were divided into two groups according to their progression after 24 months of follow-up: group 1, with overall survival of at least 24 months and group 2, deceased due to cancer in the follow-up period. The institutional research ethics committee approved this study. The demographic data and clinical information were obtained from the medical records and included: age, origin, education, smoking status, number of children, and number of sexual partners, in addition to the clinical and imaging staging data (FIGO) (4), and information regarding the treatment.

The scans were performed at the nuclear medicine and molecular imaging facility of the *Instituto de Medicina Integral Professor Fernando Figueira* using PET/CT scanner (Siemens Biography 16 channels, Germany), according to the guidelines of the European Society of Nuclear Medicine (20). Patients fasting for at least 4 h and with glycemic levels ≤ 150 mg/dL received 0.14 mCi/kg of ^{18}F -FDG intravenously. Approximately 60 min after the administration of ^{18}F -FDG, images were obtained from the skull to the thigh root. All the patients received 20 mg of furosemide after the first imaging; additionally, 120 min after the radiopharmaceutical injection, they returned to the scanner for late imaging of the pelvis. The acquisition parameters of the initial images were analyzed, with a reconstruction diameter of 500 mm, tube voltage of 130 kV, current of 75–310 mAs, and thickness of 3 mm. The images were reconstructed with 3D OSEM mode (four iterations and eight subsets) in a $4.07 \times 4.07 \times 5.00$ mm³ matrix.

TABLE 1 Clinical and demographic characteristics of the study patients.

Variable	n (%)	%
N = 47		
Age (mean ± SD)	47 ± 23 years	
Origin		
Metropolitan area	26 (53)	55.4
Inland cities	21	44.6
Education		
Illiterate	17	33.1
0 to 8 years	24	51.0
8 to 12 years	06	12.7
Smoking		
Non-smoker	23	48.9
<20 pack-year	04	8.5
>20 pack-year	13	27.6
Ex-smoker for > 5 years	08	14.8
Number of children		
1 child	6	12.7
2 children	7	14.8
3 or more children	34	72.3
Number of sexual partners		
Up to two partners	12	25.6
Three or two partners	35	74.4
Tumor size (cm) (mean ± SD)	5.43 cm (SD 1.49)	
FIGO staging		
IB2	02	4.20
II	04	8.0
III	21	44.6
IV	20	42.5

FIGO, international federation of gynecology and obstetrics.

Radiomic analysis

Segmentation

We use the free access multiplatform Local Image Features Extraction (LIFEx) software (V6.30—Inserm, Orsey, France) (21), as can be seen in **Supplementary Figure 1**. Initially, a semi-automatic segmentation of the uterine cervical lesion was performed (whole-body image only), identified by the ^{18}F -FDG uptake on the CT fusion image, and manually outlined with a 3D design tool. Subsequently, the software selected the area of highest uptake, considering a fixed threshold of 40% of the standard uptake value (SUV) of the ROI volume (VOI), a method validated for cervical uterine neoplasms (22, 23). Notably, the details regarding the computation parameters and formulas are described at www.lifexsoft.org (21). A radiologist specialized in the female pelvis and supervised by a nuclear medicine specialist, both with 20 years of experience, did the segmentations for all patients.

Extraction

For each selected volume, a massive extraction of numerical data was performed by LIFEx, using $4 \times 4 \times 4$ resizing,

0.25 fixed number width (FBW) intensity discretization method and histogram redefinition, obtaining 50 tumor features. These features were divided into categories, including: first-order statistics derived from the voxel intensity histogram (shape, volume, and histogram), and conventional indices (SUVpeak, SUVmean, SUVmax, MTV, and TLG); second-order statistics, including features based on the gray-level co-occurrence matrix (GLCM), gray-level run-length matrix (GLRLM), neighboring gray-level dependence matrix (NGLDM), and gray-level zone length matrix (GLZLM) (12, 21, 24).

Selection of radiomic features

Initially, searching for clinically significant markers associated with OS, we performed an independent sample test with the Mann–Whitney to assess the distribution for each feature in the two groups, including those with a p -value < 0.05, to subsequent analysis.

After that, the data were submitted to dimension reduction through rank correlation with Spearman's coefficient, evaluating each pair of features. Later, we found which markers correlated with each other, excluding redundant markers using a correlation matrix and selecting those with a pre-established hypothetical ρ lower than 0.85.

Statistical analysis

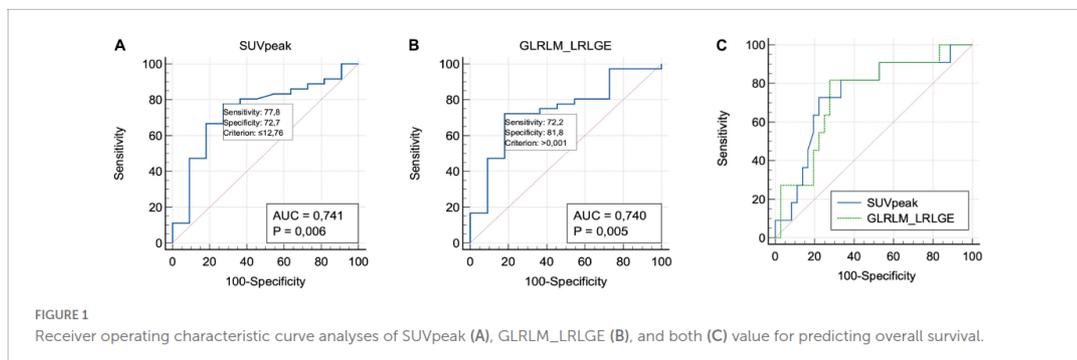
The absolute and relative frequency described categorical variables in percentage. Continuous variables with a normal distribution were analyzed by the mean and standard deviation; while non-parametric variables were analyzed by the median, maximum and minimum values, and interquartile range (IQR). For comparison between variables, we used the Mann–Whitney U test. We determined the cut-off points for variables with a p -value < 0.05 and the distinction between groups by ROC curves (DeLong methodology).

For prognostic evaluation, we correlated the selected radiomic features with the OS. Kaplan–Meier survival curves were constructed, with cut-off points obtained by the ROC curve for each variable, using the MedCalc software (MedCalc Software Ltd, Ostend, Belgium; <https://www.medcalc.org>; 2022). P -values lower than 0.05 were considered statistically significant.

Results

Clinical and demographic characteristics

The sample was initially composed of 50 consecutive patients. Three patients were excluded: one whose pretreatment baseline scan was unavailable and two other scans with divergence in the acquisition parameters (disagreement with IBSI standards).



Therefore, 47 patients were eligible for this study, with a mean age of 47 ± 23 years and ranging from 24 to 70 years. The majority of the patients had a low level of education, with reports of multiple sexual partners. Approximately 87% of patients presented with advanced stages of the disease (FIGO III and IV) (Table 1). Advanced stages of FIGO were correlated with lower overall survival (Supplementary Figure 2).

The median follow-up was 23.5 months (range: 3.73–39 months), with all surviving patients being followed up for at least 24 months. Of a total of 47 patients, 36 (77%) were alive at the end of 24 months (group 1) and 11 (23%) patients had died due to the disease (group 2).

Metabolic biomarkers and textural radiomic features

The data were extracted from 47 VOIs. We selected the features with discriminatory power for the selected outcome: three metabolic parameters—SUVmax ($p = 0.02$), SUVmean ($p = 0.02$), and SUVpeak ($p = 0.01$); and 13 textural markers—five markers from the GLZLM matrix (GLZLM_SZE, GLZLM_LGZE, GLZLM_HGZE, GLZLM_SZLGE, and GLZLM_SZHGE); six markers from the GLRLM matrix (GLRLM_LGRE, GLRLM_HGRE, GLRLM_SRLGE, GLRLM_SRHGE, GLRLM_LRLGE, and GLRLM_LRHGE); and two markers from the GLCM matrix (GLCM_Contrast variance and GLCM_Dissimilarity).

Among the metabolic parameters, the SUVpeak showed the best performance to differentiate the groups. The SUVpeak median in group 1 was 10.89 (IQR 7.60–12.69), while in group 2 was 13.87 (IQR 12.17–14.14), $p = 0.01$. The best cut-off point value (ROC curve analysis) was 12.76 with an AUC of 0.74, a sensitivity of 77.8%, and a specificity of 72.7%, $p = 0.006$.

The SUVmax median value in group 1 was 12.83 (IQR 9.09–14.90) vs. 15.98 in group 2 (IQR 13.52–19.09), $p = 0.02$. The best cut-off point was 14.32, AUC = 0.68, sensitivity = 72.3%, and specificity = 72.7% for the cut-off point of 14.32 ($p = 0.012$). The SUVmean median value in group 1 was 7.68 (IQR 9.09–14.90)

vs. 9.88 in group 2 (IQR 8.88–10.92), $p = 0.02$. It presented an AUC of 0.68, sensitivity of 72.3%, and specificity of 72.7% for a cut-off point of 8.8 ($p = 0.01$).

The other conventional metabolic metrics were not significant. The median MTV in group 1 was 31.9 (IQR: 18.5–51.0) vs. 37.8 (IQR: 24.6–72.4) in group 2 ($p = 0.49$). The median TLG in group 1 was 295.9 (IQR: 100.7–403.7) vs. 320.3 (IQR: 253.2–465.7) in group 2, $p = 0.33$.

Aiming for the redundancy feature reduction, we used the Spearman rank correlation for each of these 13 attributes. Three of them showed a rho value lower than 0.85: GLRLM_LGRE, GLRLM_SRLGE, and GLRLM_LRLGE. When we compared the AUC of these three indices, the GLRLM_LRLGE textural feature presented a little better performance. The GLRLM_LRLGE median in group 1 was 1.2×10^{-3} (IQR: 8×10^{-4} – 32×10^{-3}) vs. 7.7×10^{-3} in group 2 (IQR 6×10^{-4} – 9×10^{-3} , $p = 0.017$). The best cut-off point value was 1×10^{-3} (AUC: 0.74; sensitivity: 72.2%; specificity: 81.8%, $p = 0.005$).

For GLRLM_LGRE, the group 1 median value was 1.2×10^{-3} (IQR: 7×10^{-4} – 2.4×10^{-3}), and in group 2 was 7×10^{-4} (IQR: 5×10^{-4} to 8×10^{-4} , $p = 0.01$). It presented an AUC of 0.73, sensitivity of 81.8%, and specificity of 72.2% for a cut-off point of 9×10^{-4} ($p = 0.006$). For GLRLM_SRLGE, the group 1 median value was 1.2×10^{-3} (IQR: 7×10^{-4} – 2.3×10^{-3}), and in group 2 was 7×10^{-4} (IQR: 5×10^{-4} – 8×10^{-4} , $p = 0.01$). The AUC was 0.73, sensitivity of 81.8%, and specificity of 72.2% for a cut-off point of 8×10^{-4} ($p = 0.006$) (Figure 1). More information at Table 2.

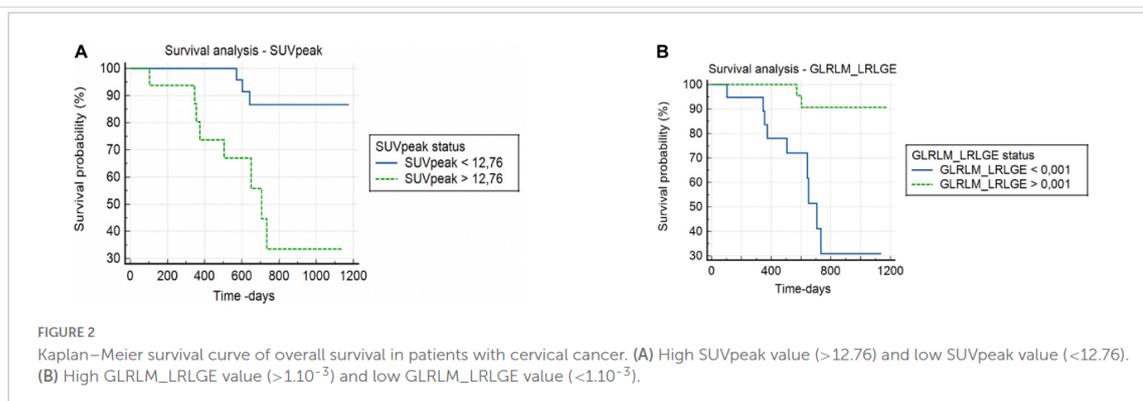
Correlations between conventional parameters ^{18}F -FDG PET/CT and textural features

The SUVpeak showed negative correlations with the GLRLM matrix. GLRLM_LRLGE ($r = -0.890$, $p < 0.01$). The SUVmax also showed a negative correlation with GLRLM_LRLGE ($r = 0.764$, $p < 0.01$).

TABLE 2 Results of independent-samples Mann–Whitney test analysis parameters 18F-FDG PET/CT for cervical cancer predicting overall survival.

	Group 1		Group 2		P
	Median	Range	Median	Range	
Image-based parameters					
SUV mean	7.68	9.09–14.90	9.88	8.88–10.92	0.02
SUV peak	10.89	7.60–12.69	13.87	12.17–14.14	0.01
SUV max	12.83	9.09–14.90	15.98	13.52–19.09	0.02
TLG (mL)	295.9	100.7–403.7	320.3	253.2–465.7	0.33
MTV (mL)	31.9	18.5–51.0	31.9	24.6–72.4	0.49
Texture parameters					
GLRLM_LGRE	1.2×10^{-3}	7×10^{-4} – 2.4×10^{-3}	7×10^{-4}	5×10^{-4} – 8×10^{-4}	0.01
GLRLM_SRLGE	1.2×10^{-3}	7×10^{-4} – 2.3×10^{-3}	7×10^{-4}	5×10^{-4} – 8×10^{-4}	0.01
GLRLM_LRLGE	1.2×10^{-3}	8×10^{-4} – 3.2×10^{-3}	7.7×10^{-3}	6×10^{-4} – 9×10^{-4}	0.01

¹⁸F-FDG PET/CT, ¹⁸fluorodeoxyglucose positron emission tomography; Group 1, survivors; Group 2, died; SUV, standardized uptake values; SUVmax, maximum standardized uptake value; SUVmean, mean standardized uptake value; SUVpeak, the peak of SUV in 1 mL; TLG, total lesion glycolysis; MTV, metabolic tumor volume; GLRLM, Gray level run length matrix; LGRE, low gray-level runs emphasis; SRLGE, short runs low gray-level emphasis; LRLGE, long runs gray-level emphasis.



Both SUVpeak and GLRLM_LRLGE were not correlated with FIGO staging (Supplementary Table 1).

Kaplan–Meier survival analysis

GLRLM_LRLGE showed a significant correlation with the OS ($p = 0.003$). Patients who died presented a GLRLM_LRLGE value lower than the cut-off point, with a shorter survival time: median of 708 days (CI: 505.0–734.0). A risk ratio of 10.8 (CI: 3.0–39.1) was observed.

SUVpeak showed a significant correlation with the OS ($p = 0.006$). Patients who died presented a higher SUVpeak value, with a shorter survival time: a median of 706 days (CI: 374.0–734.0). A risk ratio of 10.5 (CI: 2.7–40.3) was observed (Figure 2).

Discussion

This study demonstrated the prognostic association between radiomic biomarkers of primary uterine cervical cancer lesions

at ¹⁸F-FDG-PET/CT and overall survival. Among the evaluated metabolic parameters, SUVpeak showed the best discriminatory power; and among all the selected radiomic textural features, the GLRLM_LRLGE presented the best predictive performance. Moreover, SUVpeak and GLRLM_LRLGE demonstrated a greater correlation with OS compared with clinical and other more conventional ¹⁸F-FDG-PET/CT parameters, including MTV and TLG. These data reinforce the importance of metabolic radiomic evaluation in cervical uterine tumor staging.

Cervical cancer accounts for high morbidity and mortality in patients of productive and reproductive age worldwide, especially in vulnerable populations (25). The staging of this neoplasm is based on FIGO classification, which includes characteristics of the primary lesion, and lymph node or distant dissemination (4–6). However, FIGO classification presented a low accuracy in predicting therapy response and survival, especially among patients with advanced-stage cancer disease (4).

The search for non-invasive and robust prognostic biomarkers can improve the predictive power of therapy response. Radiomic is considered a promising analysis tool in

precision medicine (26, 27). Some studies have also reported the use of this technology in cervical tumor cases, based on several imaging methods, especially magnetic resonance imaging and ^{18}F -FDG-PET/CT (28). Usually, these studies aim to evaluate several aspects of the tumor, covering most frequently lymph node invasion (29–31), prognosis (28, 32), and therapeutic response (33); followed by histological grade (34–36), staging (37), and lymphovascular space invasion (29).

Standardized uptake value represents a semi-quantitative metric of ^{18}F -FDG-PET/CT with prognostic ability, including OS evaluation of patients with uterine cervix tumors (10, 33). All metrics of SUV are correlated with each other, providing information on the tumor metabolic activity. SUVpeak is reported as more robust and reproducible than SUVmax and SUVmean, although it is not widely disseminated in clinical practice (7). Studies report better performance of SUVpeak to demonstrate the aggressiveness of early-stage cervical tumors compared to SUVmax, maybe because SUVpeak measures several voxels in a more metabolically active spherical VOI of the lesion (7, 38).

The SUVpeak in our study presents a cut-off value similar to those described in other studies. Schernberg et al. (39) analyzed locally advanced disease treated with definitive chemoradiation and demonstrated that a high SUVpeak value was superior in predicting the OS and local recurrence, when compared with other ^{18}F -FDG-PET/CT parameters, like MTV and TLG. Other studies also evaluated early-stage cervical cancer, in which a low SUVpeak was significantly correlated with high progression-free survival (40).

A systematic review by Piñeiro-Fiel et al. evaluated the radiomics of ^{18}F -FDG-PET/CT in several neoplasms. Gynecological cancers were among the four most studied types, with 19 publications in a total of 741 studies. Of these 19 publications, cervical uterine cancer accounted for the largest number of publications (74%), followed by endometrial cancer (16%). As in our study, the textural features were correlated with the conventional metrics of ^{18}F -FDG-PET/CT, including SUV. In the analysis of gynecological cancers, the texture matrices that presented higher significance were GLCM, GLRLM, and GLZSM (41).

We showed that GLRLM_LRLGE could perform well in predicting OS in patients with advanced cervical cancer. The radiomic matrix GLRLM conceptually relates to the intensity of the gray level of pixels in an image, in a given direction, and LRLGE represents the distribution of long stretches with a high or low gray level, being an indicator of the uniformity of the homogeneous distribution of FDG uptake (42). GLRLM_LRLGE is a potential biomarker in other neoplasms too, as it can discriminate benign from malignant renal tumors (42), and can be used to assess recurrence in rectal cancer (43).

Additionally, some studies demonstrated a significant correlation between the GLRLM matrix ^{18}F -FDG-PET/CT

textural markers (LGRE, SRLGE, and LRLGE) and RNA-level immunological biomarkers of PD-L1 (programmed death ligand 1) in lung cancer (44). PD-L1 protein expression is also a predictive biomarker in uterine cervical cancer (45). Subsequently, we could find a possible intercorrelation between these textural markers (GLRLM_LRLGE) and PD-L1 expression, representing an important prognostic and selection factor for immunotherapy. This hypothesis may be evaluated in future prospective studies.

GLRLM_LRLGE possibly shows a relationship with tumor necrosis, as it assesses the homogeneity of ^{18}F -FDG uptake, and its highest value is documented in benign homogeneous lesions (42). On the other hand, several studies demonstrate a direct relationship between PD-L1 and tumor necrosis factor (TNF alpha) in oncologic diseases, including findings in which TNF alpha produced by adipocytes positively regulates PD-L1 (46). Based on these findings, we can assume that the textural factor GLRLM_LRLGE also correlates with TNF alpha.

However, this study has many limitations. It is a single-center study with a low number of patients and a retrospective analysis. However, the sample was derived from a clinical trial (47), with a relatively homogeneous and controlled group of patients with a good clinical follow-up. Additionally, the ^{18}F -FDG-PET/CT pretreatment images were reevaluated in order to collect new data regarding radiomic characteristics in the primary lesions. Moreover, we analyzed only the scans with a protocol following the parameters established by the IBSI. We also do not perform multiple correction tests in our data, mainly because of the low number of included patients.

In conclusion, in patients with advanced cervical tumors, this study investigated and identified two biomarkers with better prognostic performance (SUVpeak and GLRLM_LRLGE). These features denote metabolism and intratumoral textural homogeneity, respectively. In the future, the SUVpeak and GLRLM_LRLGE have the potential to be incorporated into clinical practice, helping to identify patients with a higher risk of death.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Comitê de Ética em Pesquisa do IMIP. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

NA, SB, and MM conceptualized and designed the study. T-MC and CA were acquired the data. NA and TM identified and marked the lesions. MM and FM inferred results, implemented the methods, and analyzed and interpreted the data. NA, MO, and SB drafted the manuscript and wrote the first draft of the manuscript. LM, FM, SB, and MO critically revised the manuscript and provided supervision, support, conceptualization, and guidance throughout the project. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships

that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmed.2022.1046551/full#supplementary-material>

References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* (2018) 68:394–424. doi: 10.3322/caac.21492
- Arbyn M, Weiderpass E, Bruni L, Sanjosé S, Saraiya M, Ferlay J, et al. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *Lancet Glob Health.* (2019) 8:e191–120. doi: 10.1016/S2214-109X(19)30482-6
- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* (2021) 71:209–49. doi: 10.3322/caac.21660
- Bhatla N, Denny L. FIGO Cancer Report 2018. *Int J Gynaecol Obstet.* (2018) 143(Suppl. 2):2–3. doi: 10.1002/ijgo.12608
- Lee SI, Atri M. 2018 FIGO Staging system for uterine cervical cancer: enter cross-sectional imaging. *Radiology.* (2019) 292:15–24. doi: 10.1148/radiol.2019190088
- Salib MY, Russell JHB, Stewart VR, Sudderuddin SA, Barwick TD, Rockall AG, et al. 2018 FIGO staging classification for cervical cancer: added benefits of imaging. *Radiographics.* (2020) 40:1807–22. doi: 10.1148/rg.2020200013
- Sher A, Lacoueille F, Fosse P, Vervueren L, Cahouet-Vannier A, Dabli D, et al. For avid glucose tumors, the SUV peak is the most reliable parameter for [(18)F]FDG-PET/CT quantification, regardless of acquisition time. *EJNMMI Res.* (2016) 6:21. doi: 10.1186/s13550-016-0177-8
- Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, et al. European Association of Nuclear Medicine (EANM). FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging.* (2015) 42:328–54. doi: 10.1007/s00259-014-2961-x
- Im HJ, Bradshaw T, Solaiyappan M, Cho SY. Current methods to define metabolic tumor volume in positron emission tomography: which one is better? *Nucl Med Mol Imaging.* (2018) 52:5–15. doi: 10.1007/s13139-017-0493-6
- Herrera FG, Breuneval T, Prior JO, Bourhis J, Ozsahin M. 18F-FDG-PET/CT metabolic parameters as useful prognostic factors in cervical cancer patients treated with chemo-radiotherapy. *Rad Oncol.* (2016) 11:1–11. doi: 10.1186/s13014-016-0014-x
- Bollineni VR, Ytre-Hauge S, Gulati A, Halle MK, Woie K, Salvesen O. The prognostic value of preoperative FDG-PET/CT metabolic parameters in cervical cancer patients. *Eur J Hybrid Imaging.* (2018) 2:1–14. doi: 10.1186/s41824-018-0042-2
- Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology.* (2016) 278:563–77. doi: 10.1148/radiol.201511169
- van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* (2017) 77:e104–7. doi: 10.1158/0008-5472
- Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and challenges of image analysis. *Eur Radiol Exp.* (2018) 2:36. doi: 10.1186/s41747-018-0068-z
- Anbumani S, Jayaraman P, Anchaneyan P, Bilimappa RS, Nambiraj A. Quantitative radiomic phenotyping of cervix cancer. *Int Clin Pathol J.* (2018) 6:26–8. doi: 10.15406/icpj.2018.06.00149
- Cook GJR, Siddique M, Taylor BP, Yip C, Chicklore S, Goh V. Radiomics in PET: principles and applications. *Clin Transl Imaging.* (2014) 2:269–76. doi: 10.1007/s40336-014-0064-0
- Ho KC, Fang YHD, Chung HW, Yen TC, Ho TY, Chou HH, et al. A preliminary investigation into textural features of intratumoral metabolic heterogeneity in (18)F-FDG PET for overall survival prognosis in patients with bulky cervical cancer treated with definitive concurrent chemo-radiotherapy. *Am J Nucl Med Mol Imaging.* (2016) 6:166–75.
- Mu W, Liang Y, Hall LO, Tan Y, Balagurunathan Y, Wenham R, et al. 18F-FDG PET/CT Habitat radiomics predicts outcome of patients with cervical cancer treated with chemo-radiotherapy. *Radiol Artif Intell.* (2020) 2:e190218. doi: 10.1148/ryai.2020190218
- Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative. *arXiv.* (2016) [Preprint]. doi: 10.48550/arXiv.1612.07003
- Boellaard R, O'Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Stroobants SG, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging.* (2010) 37:181–200. doi: 10.1007/s00259-009-1297-4
- Nioche C, Orlhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, et al. LIFEX: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res.* (2018) 78:4786–9. doi: 10.1158/0008-5472
- Yang F, Thomas MA, Dehdashti F, Grigsby PW. Temporal analysis of intratumoral metabolic heterogeneity characterized by textural features in cervical cancer. *Eur J Nucl Med Mol Imaging.* (2013) 40:716–27. doi: 10.1007/s00259-012-2332-4

23. Tsujikawa T, Rahman T, Yamamoto M, Yamada S, Tsuyoshi H, Kiyono Y, et al. 18F-FDG PET radiomics approaches: comparing and clustering features in cervical cancer. *Ann Nucl Med.* (2017) 31:678–85. doi: 10.1007/s12149-017-1199-7
24. Thibault G, Fertil B, Navarro C, Pereira S, Cau P, Levy N, et al. Texture indexes and gray level size zone matrix application to cell nuclei classification. *Proceedings of the 10th International Conference on Pattern Recognition and Information Processing.* Minsk (2009).
25. Stelzle D, Tanaka LF, Lee KK, Khalil AI, Baussano I, Shah ASV, et al. Estimates of the global burden of cervical cancer associated with HIV. *Lancet Glob Health.* (2021) 9:e161–9. doi: 10.1016/S2214-109X(20)30459-9
26. Barbet J, Bernaudin M, Payoux P, Cicone F, Gaugler MH, Kraeber-Bodéré F. Editorial: nuclear medicine in the context of personalized medicine. *Front Med.* (2020) 7:252. doi: 10.3389/fmed.2020.00252
27. Ai Y, Zhu H, Xie C, Jin X. Radiomics in cervical cancer: current applications and future potential. *Crit Rev Oncol Hematol.* (2020) 152:102985. doi: 10.1016/j.critrevonc.2020.102985
28. Lucia F, Visvikis D, Vallières M, Desseroit MC, Miranda O, Robin P, et al. External validation of a combined PET and MRI radiomics model for prediction of recurrence in cervical cancer patients treated with chemo-radiotherapy. *Eur J Nucl Med Mol Imaging.* (2019) 46:864–77. doi: 10.1007/s00259-018-4231-9
29. Li Z, Li H, Wang S, Dong D, Yin F, Chen A, et al. MR-based radiomics nomogram of cervical cancer in prediction of the lymph-vascular space invasion preoperatively. *J Magn Reson Imaging.* (2019) 49:1420–6. doi: 10.1002/jmri.26531
30. Chen X, Liu W, Thai TC, Castellano T, Gunderson CC, Moore K, et al. Developing a new radiomics-based CT image marker to detect lymph node metastasis among cervical cancer patients. *Comput Methods Programs Biomed.* (2020) 197:105759. doi: 10.1016/j.cmpb.2020.105759
31. Song J, Hu Q, Ma Z, Zhao M, Chen T, Shi H. Feasibility of TWI-MRI-based radiomics nomogram for predicting normal-sized pelvic lymph node metastasis in cervical cancer patients. *Eur Radiol.* (2021) 31:6938–48. doi: 10.1007/s00330-021-07735-x
32. Ferreira M, Lovinfosse P, Hermesse J, Decuypere M, Rousseau C, Lucia F, et al. [F]FDG PET radiomics to predict disease-free survival in cervical cancer: a multi-scanner/center study with external validation. *Eur J Nucl Med Mol Imaging.* (2021) 48:3432–43. doi: 10.1007/s00259-021-05303-5
33. Reuzé S, Orlhac F, Chargari C, Nioche C. Prediction of cervical cancer recurrence using textural features extracted from 18F-FDG PET images acquired with different scanners. *Oncotarget.* (2017) 8:43169–79. doi: 10.18632/oncotarget.17856
34. Shen WC, Chen SW, Liang JA, Hsieh TC, Yen KY, Kao CH. [18F]Fluorodeoxyglucose positron emission tomography for the textural features of cervical cancer associated with lymph node metastasis and histological type. *Eur J Nucl Med Mol Imaging.* (2017) 44:1721–31. doi: 10.1007/s00259-017-3697-1
35. Liu Y, Zhang Y, Cheng R, Liu S, Qu F, Yin X, et al. Radiomics analysis of apparent diffusion coefficient in cervical cancer: a preliminary study on histological grade evaluation. *J Magn Reson Imaging.* (2019) 49:280–90. doi: 10.1002/jmri.26192
36. Su X, Chen N, Sun H, Liu Y, Yang X, Wang W, et al. Automated machine learning based on radiomics features predicts H3 K27M mutation in midline gliomas of the brain. *Neuro Oncol.* (2020) 22:393–401. doi: 10.1093/neuonc/noz184
37. Umutlu L, Nensa F, Demircioglu A, Antoch G, Herrmann K, Forsting M, et al. Radiomics Analysis of Multiparametric PET/MRI for N- and M-Staging in patients with primary cervical cancer. *Rofo.* (2020) 192:754–63. doi: 10.1055/a-1100-0127
38. Vanderhoek M, Perlman SB, Jerai R. Impact of the definition of peak standardized uptake value on quantification of treatment response. *J Nucl Med.* (2012) 53:4–11. doi: 10.2967/jnumed.111.093443
39. Schernberg A, Reuze S, Orlhac F, Buvat I, Dercle L, Sun R, et al. A score combining baseline neutrophilia and primary tumor SUV_{peak} measured from FDG PET is associated with outcome in locally advanced cervical cancer. *Eur J Nucl Med Mol Imaging.* (2018) 45:187–95. doi: 10.1007/s00259-017-3824-z
40. Zhang L, Sun H, Du S, Xu W, Xin J, Guo Q. Evaluation of 18F-FDG PET/CT parameters for reflection of aggressiveness and prediction of prognosis in early-stage cervical cancer. *Nucl Med Commun.* (2018) 39:1045–52. doi: 10.1097/MNM.0000000000000909
41. Piñeiro-Fiel M, Moscoso A, Pubul V, Ruibal Á, Silva-Rodríguez J, Aguiar P. A systematic review of PET textural analysis and radiomics in cancer. *Diagnostics.* (2021) 11:380. doi: 10.3390/diagnostics11020380
42. Matsumoto S, Arita Y, Yoshida S, Fukushima H, Kimura K, Yamada I, et al. Utility of radiomics features of diffusion-weighted magnetic resonance imaging for differentiation of fat-poor angiomyolipoma from clear cell renal cell carcinoma: model development and external validation. *Abdom Radiol.* (2022) 47:2178–86. doi: 10.1007/s00261-022-03486-5
43. Park H, Kim KA, Jung JH, Rhie J, Choi SY. MRI features and texture analysis for the early prediction of therapeutic response to neoadjuvant chemoradiotherapy and tumor recurrence of locally advanced rectal cancer. *Eur Radiol.* (2020) 30:4201–11. doi: 10.1007/s00330-020-06835-4
44. Kim BS, Kang J, Jun S, Kim H, Pak K, Kim GH, et al. Association between immunotherapy biomarkers and glucose metabolism from F-18 FDG PET. *Eur Rev Med Pharmacol Sci.* (2020) 24:8288–95. doi: 10.26355/eurrev_202008_22625
45. Rotman J, den Otter LAS, Bleeker MCG, Samuels SS, Heeren AM, Roemer MGM, et al. PD-L1 and PD-L2 expression in cervical cancer: regulation and biomarker potential. *Front Immunol.* (2020) 17:596825. doi: 10.3389/fimmu.2020.596825
46. Li Z, Zhang C, Du JX, Zhao J, Shi MT, Jin MW. Adipocytes promote tumor progression and induce PD-L1 expression via TNF- α /IL-6 signaling. *Cancer Cell Int.* (2020) 20:179. doi: 10.1186/s12935-020-01269-w
47. Azevedo CRAS, Thuler LCS, Mello MJG, Lima JTO, Fonte ALF, Fontão DFS, et al. Phase II trial of neoadjuvant chemotherapy followed by chemoradiation in locally advanced cervical cancer. *Gynecol Oncol.* (2017) 146:560–5. doi: 10.1016/j.ygyno.2017.07.006

Artigo nº 6

Multi-center Integrating Radiomics, Structured Reports and Machine Learning Algorithms for Assisted Classification of COVID-19 in Lung Computed Tomography

Journal of Medical and Biological Engineering

Submetido

**Multi-center Integrating Radiomics, Structured Reports and Machine Learning
Algorithms for Assisted Classification of COVID-19 in Lung Computed Tomography**

Marcos A. D. Machado

*Department of Radiology
Complexo Hospitalar Universitário Prof. Edgard Santos/ Ebserh
Universidade Federal da Bahia
Salvador, Bahia
Zip code: 40110-040
Brazil*

*Radtec Serviços em Física Médica
Salvador, Bahia
Zip code: 40060-330
Brazil*

ORCID <https://orcid.org/0000-0002-0106-9769>

Ronnyldo R. E. Silva

*Department of Systems and Computing
Universidade Federal de Campina Grande
Campina Grande, Paraíba
Zip code: 58429-900
Brazil*

*Radtec Serviços em Física Médica
Salvador, Bahia
Zip code: 40060-330
Brazil*

ORCID <https://orcid.org/0000-0002-4594-0886>

Mauro Namias

*Department of Medical Physics
Nuclear Diagnostic Center Foundation
Buenos Aires
Zip code: C1417CVE
Argentina*

ORCID <https://orcid.org/0000-0001-5808-4102>

Andreia S. Lessa

*Department of Radiology
Hospital Universitário Gaffrée e Guinle
Universidade do Rio de Janeiro (UNIRIO)
Rio de Janeiro, Rio de Janeiro
Zip code: 20270-004
Brazil*

ORCID <https://orcid.org/0000-0002-1347-4868>

Margarida C. L. C. Neves

*Department of Pneumology
Complexo Hospitalar Universitário Prof. Edgard Santos/ Ebserh
Universidade Federal da Bahia
Salvador, Bahia
Zip code: 40110-040
Brazil*

ORCID <https://orcid.org/0000-0002-2537-3079>

Carolina T. A. Silva

*Department of Radiology
Complexo Hospitalar Universitário Prof. Edgard Santos/ Ebserh
Universidade Federal da Bahia
Salvador, Bahia
Zip code: 40110-040
Brazil*

ORCID <https://orcid.org/0000-0001-5489-0801>

Danillo M. Oliveira

*Department of Radiology
Hospital Universitário Alcides Carneiro/ Ebserh
Universidade Federal de Campina Grande
Campina Grande, Paraíba
Zip code: 58400-398
Brazil*

*Northeast Regional Nuclear Science Centre (CRCN-NE)
Recife, Pernambuco
Zip Code: 50840-545
Brazil*

*Nuclear Energy Department
Universidade Federal de Pernambuco*

*Recife, Pernambuco
Zip code: 50740-540
Brazil*

ORCID <https://orcid.org/0000-0003-3540-7663>

Thamiris R. Reina

*Department of Radiology
Hospital Universitário da Universidade Federal de Juiz de Fora/ Ebserh
Universidade Federal de Juiz de Fora
Juiz de Fora, Minas Gerais
Zip code: 36038-330
Brazil*

ORCID <https://orcid.org/0000-0002-4012-0352>

Arquimedes A. B. Lira

*Department of Radiology
Hospital Universitário Alcides Carneiro/ Ebserh
Universidade Federal de Campina Grande
Campina Grande, Paraíba
Zip code: 58400-398
Brazil*

ORCID <https://orcid.org/0000-0002-3553-5254>

Leandro M. Almeida

*Centro de Informática
Universidade Federal de Pernambuco
Recife, Pernambuco
Zip Code: 50720-001
Brazil*

ORCID <https://orcid.org/0000-0001-8025-0517>

Cleber Zanchettin

*Centro de Informática
Universidade Federal de Pernambuco
Recife, Pernambuco
Zip Code: 50720-001
Brazil*

Department of Chemical and Biological Engineering

*Northwestern University
Evanston, IL 60208, USA*

ORCID <https://orcid.org/0000-0001-6421-9747>

Eduardo M. Netto

*Infectious Disease Research Laboratory
Complexo Hospitalar Universitário Prof. Edgard Santos/ Ebserh
Universidade Federal da Bahia
Salvador, Bahia
Zip code: 40110-040
Brazil*

ORCID <https://orcid.org/0000-0003-1691-6761>

Correspondence to: Marcos Machado, *E-mail*: machado@radtec.com.br

running head: Artificial intelligence for assisted CT COVID-19 classification

Abbreviations:

API: application programming interface

AI: artificial intelligence

AUC: area under ROC curve

CI: 95% confidence interval

COVID-19: coronavirus disease 2019

CT: computed tomography

ICC: intraclass correlation coefficient

ML: machine learning

PACS: Picture Archiving Communication System

RIS: Radiology Information System

ROC: receiver operating characteristics

xgboost: extreme gradient boosting

ABSTRACT

Purpose: To evaluate the classification performance of structured report features, radiomics, and machine learning models to differentiate COVID-19 from other types of pneumonia in chest CT images. **Methods:** Sixty-four COVID-19 subjects and 64 subjects with non-COVID-19 pneumonia were sequentially selected. The dataset was split into two independent cohorts, one for the structured report, the radiomic feature selection and model building (n=73), and another for model validation (n=55). Physicians performed readings with and without machine learning support. Model's sensitivity and specificity were calculated, and inter-rater reliability was assessed using Cohen's Kappa agreement. **Results:** Physicians performed with mean sensitivity and specificity of 83.4% and 64.3%, respectively. When they were assisted with machine learning, mean sensitivity and specificity were 87.1% and 91.1%, respectively. In addition, machine learning improved the inter-rater reliability from moderate to substantial agreement. **Conclusions:** Integrating structured reports and radiomics promises assisted the classification of COVID-19 in computed tomography images of the chest.

Keywords: radiomics, structured reports, machine learning, pneumonia, computed tomography.

1. INTRODUCTION

Computed tomography holds a significant place in the diagnosis and staging of pneumonia, therefore providing important information in managing the current COVID-19 pandemic [1]. However, the aspects of pulmonary lesions seen in CT are typical but not pathognomonic [2], resulting in a limitation of the method [3].

Artificial intelligence (AI) is a ubiquitous topic in medical imaging due to its ability to automate image segmentation, improve lesion detection, perform better quality assurance, and

increase the inter- and intra-rater reliability [4-7]. Furthermore, structured reports are claimed to enhance radiology reporting by enabling standardization, productivity, and improved information transmission [8,9], leveraging its potential to decision support technologies using AI [8,10,11].

Recent studies have shown promising radiomics and AI to classify pneumonia in CT studies [12,13]. Even though AI has been widely used with different data types for COVID-19 diagnoses, such as radiomics and clinical data, it has been little explored using imaging reports [14]. Huanhuan Liu and colleagues developed a clinical-radiomics model that included distribution features and clinical data, demonstrating the added value of lesion distribution reporting in their models [13]. In this perspective, the integration of minable data with artificial intelligence may benefit radiology practices and patient care.

In this work, we explored radiological features from multi-center standardized web-based structured reports and radiomics. In addition, we developed AI models and evaluated their classification performance to differentiate COVID-19 from other types of pneumonia, and their impact on intra rater reliability.

2. MATERIALS AND METHODS

2.1 Subjects

Cases were defined as subjects with COVID-19 confirmed with RT-PCR referred to thorax CT during April-2020 and April-2021 sequentially selected from three Brazilian academic hospitals (Table 1). Controls were defined as subjects sequentially selected with pneumonia before the COVID-19 pandemic, referred to thorax CT during January-2018 and October-2019.

Exclusion criteria were (a) CT images with poor quality; (b) small or imperceptible lesions; (c) unavailability of RT-PCT for the confirmed COVID-19 group; (d) images that were

not compliant with the institutional protocol for lung imaging (e.g., matrix size, kernel), and e) images that were no longer available on the PACS at the retrieval.

Table 1 – Subject and center characteristics

Characteristic	Hospital A	Hospital B	Hospital C
Scanner	Brilliance 6 (Philips)	Aquilion (Toshiba)	Emotion 6 (Siemens)
COVID/ non-COVID	25/ 27	25/ 20	14/ 17
Lung involvement			
Mild (<25%)	18	26	11
Moderate (25-50%)	20	7	11
Severe (>50%)	14	12	9
Male gender	25	18	12
Age (median – IQR)	54.5 (43.0-70.8)	55.5 (41.3-64.8)	57.0 (51.5-66.5)
Slice thickness	2.0 mm	2.0 mm	1.25 mm
Energy (kV)	120	120	130
Current (mA)	159-310	99–389	61-240
Kernel	L	FC86	B30s
Experience (years)	11	10	19

M: male; IQR: interquartile range. Experience: period of radiologist experience with CT

The study was approved by the ethics committee at each institution: Hospital A, Hospital B and Hospital C.

2.2 Structured report

The structured report was composed of ten questions concerning the radiological findings: five about lesion analysis and five about lesion distribution, with each answer being binary (yes/no), as follows:

Analysis:

A1: Predominant ground-glass opacity with a rounded morphology

A2: ground-glass opacities with superimposed interlobular septal thickening and intralobular septal thickening (crazy paving)

A3: Ground glass opacity and pulmonary consolidation

A4: Pulmonary consolidation with air bronchograms

A5: Reversed halo sign or cryptogenic organizing pneumonia (COP) findings

Distribution:

D1: Bilateral and multifocal

D2: Peripheral distribution

D3: Prevalent in lower lobes and dorsal region

D4: Peribronchovascular opacities and peripheral distribution.

D5: Diffuse opacities

Experienced radiologists (one from each hospital) were asked to retrieve the corresponding subject CT images from the hospital PACS and select the findings for the structured report. Only the image series with the axial slices of the lung were allowed to be reviewed. This data was then included in the software Nuclearis (Salvador, Brazil), a web-

based radiology information system capable of personalizing standard structured reports. No other information was analyzed (e.g., side notes, remarks, further information).

2.2.1 Structured report model

An Extreme Gradient Boosting (XgBoost) algorithm was built in the training cohort (n=73) and validated in the independent cohort (n=55) using a structured report score of 0 or 1, with 0 meaning non-COVID-19 and 1 meaning COVID-19.

2.3 Radiomics

2.3.1 Automatic segmentation

The process involved 37 COVID-19 individuals and another 36 individuals with pneumonia not associated with COVID-19 (hereafter denominated feature selection cohort). These CT images were previously manually segmented using LIFEx, version 6.30 (www.lifex.org), to train a 2D convolutional neural network [15] with a 5-fold cross-validation scheme. A total of 12,780 CT slices were used, and data were augmented using 1 mm erosion in the LIFEx ROIs, flip, and rotation in the original images and ROIs, resulting in 51,120 images. Each training dataset used 60 subjects to segment the additional subjects automatically. The output 2D class probability maps were filtered with a 3D Gaussian kernel before thresholding to obtain the final 3D lesion segmentation. Only the independently segmented lesions were included in the radiomic feature selection analysis. This process used the established image segmentation software LIFEx to delineate the ground truth of lung lesions aiming to develop our own automatic segmentation software for the radiomic feature extraction.

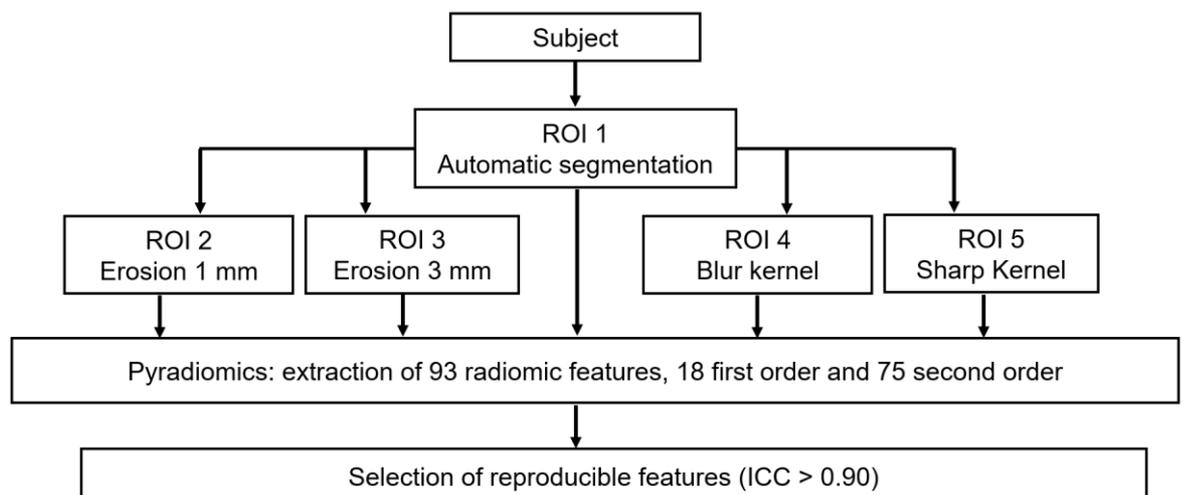
2.3.2 Feature extraction

Radiomic features were extracted using PyRadiomics [16]. The matrix grid was resampled to voxels of $1 \times 1 \times 2 \text{ mm}^3$, and the discretization within each ROI was scaled to 128 grayscale levels. Radiomic feature classes and corresponding features are presented in the supplementary material. A total of 93 features were calculated, including 18 first order and 75 textural features: 16 *gray level run length matrix* (GLRLM), 16 *gray level size zone matrix* (GLSZM), five *neighborhood gray-tone difference matrix* (NGTDM), 14 *gray level dependence matrix* (GLDM) and 24 *gray level co-occurrence matrix* (GLCM).

2.3.3 Feature selection

The feature selection process is illustrated in Figure 1. Lung lesions from the feature selection cohort were automatically segmented, and two additional segmentation datasets (artificial observers) were generated using erosion functions with 1- and 3-mm. Additional blur and sharp kernels were applied to the original image to simulate different scanner/image reconstructions. Ninety-three radiomic features were computed for each of the three observers from the erosion and image quality groups. The intraclass correlation coefficient (ICC) was used to estimate the reproducibility. Features which $\text{ICC} > 0.90$ were selected.

Figure 1: Feature selection framework



2.5.2 Radiomic model

We used the same cohort of feature selection section for model building (n=73) and an independent cohort with 55 subjects (27 COVID-19 and 28 non-COVID-19) for model validation. A random forest model was built at the lesion level (segmented lesions = 1060), each annotated according to the subject status. The algorithm classified lesions through a radiomic score ranging from 0 to 1, with 0 meaning non-COVID-19 and 1 meaning COVID-19 lesion. The model (rf_mean) was validated in the subject level by averaging the scores of all lesions for the specific subject.

2.4 Physician's performance

Two experienced physicians, one radiologist, and one pneumologist scored the CT images in the validation group (n=55) to classify whether the pneumonia was COVID-19 or non-COVID-19. Both physicians were blinded regarding the RT-PCR and clinical information, where the radiologist was the same performing the structured report (Hospital B), and the pneumologist only participated in this validation step of the study. Firstly, they analyzed the CT images without any support from the machine learning models. Finally, they analyzed the CT images again, after 30 days, with the support of both radiomic and structured reports models using binary classification in each model.

2.5 Statistical analysis

The odds ratios of radiological findings were assessed as an indication of feature importance. The whole dataset (n=128) was used in such analysis. The mean of Dice Coefficient (F1 Score) was computed using Python-3.9.2 software to assess the performance of our automatic COVID-19 segmentation method. Machine learning analysis was performed using software Orange, version 3.29.3 [17]. The sensitivity and specificity of the models were calculated from the categorical classification

outputs from Orange, and their 95% confidence intervals were estimated using OpenEpi software [18]. The ROC curves were assessed to estimate the AUC and the 95% confidence intervals using the continuous scores from orange and the SPSS 18.0 for Windows. The mean ROC curve for physicians' performance was simulated by iteratively misspecifying the true values with a random gaussian function for improved readability and interpretation. The procedure was stopped when the simulated ROC curve yielded values corresponding to the real sensitivity and specificity [19]. The inter-rater reliability was assessed with and without the support of machine learning models using the Cohen's Kappa coefficient (k), according to the classification: weak ($k \leq 0.20$), fair ($0.20 < k \leq 0.40$), moderate ($0.40 < k \leq 0.60$), substantial ($0.60 < k \leq 0.80$) or strong ($k > 0.80$) [20].

3. RESULTS

3.1 Structure reports

The XgBoost performed with 81.6% (66.6-90.8%) sensitivity and 82.9% (67.3-91.9%) specificity in the training cohort (AUC=0.91); and 77.3% (56.6-89.9%) sensitivity and 69.7% (52.7-82.7%) in the validation cohort (AUC=0.82) (Table 3).

The odds ratio (95% CI) revealed feature importance, as presented in Table 2. According to this analysis, features A1, A2, A4, D1, and D2 were the strongest predictors ($p < 0.05$), where their presence was more frequent in the COVID-19 group, except A4, which appear more frequently in the non-COVID-19 group.

Table 2 – Feature importance for structured reports

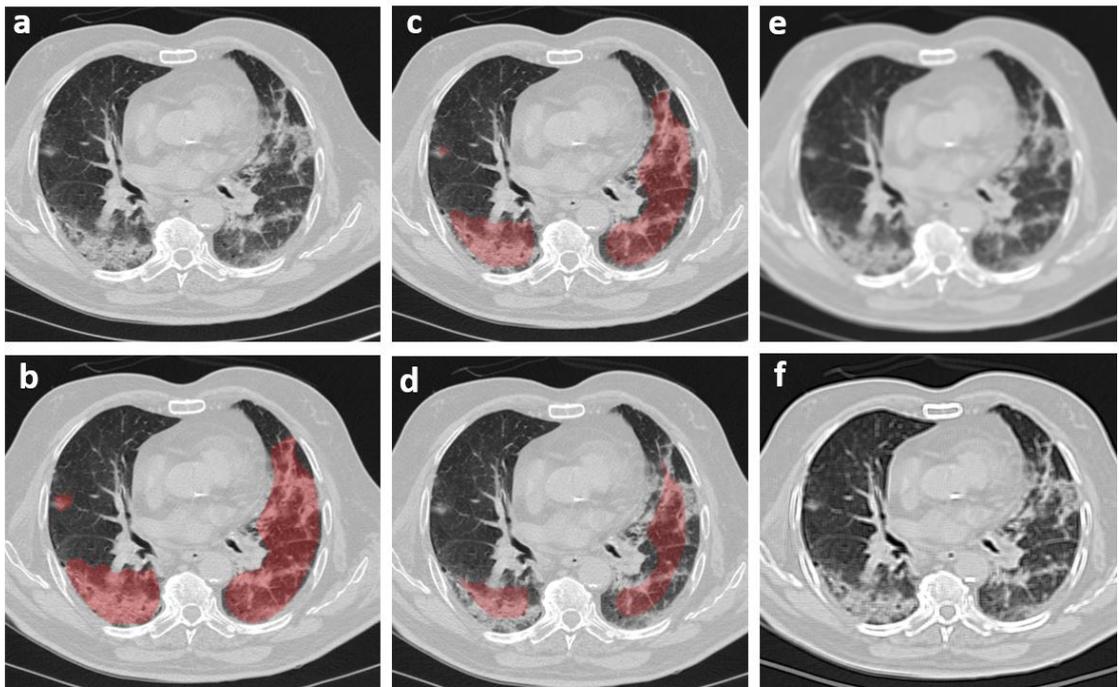
Odds Ratio (95% CI)	
Analysis	Distribution
A1: 2.7 (1.3-5.6)	D1: 2.7 (1.4-5.4)
A2: 2.6 (1.3-5.3)	D2: 5.0 (2.2-11.2)

A3: 1.4 (0.7-2.8)	D3: 0.6 (0.3-1.3)
A4: 0.3 (0.1-0.8)	D4: 1.8 (0.9-3.7)
A5: 1.0 (0.1-15.9)	D5: 1.5 (0.5-4.0)

3.2 Radiomics

The mean F1 score of our auto-segmentation algorithm was 0.72. Figure 2 shows the axial view of an infected COVID-19 individual (Figure 2a), the auto-segmentation (Figure 2b), ROI erosion functions: 1mm (Figure 2c) and 3mm (Figure 2d); and image enhancements: blur (Figure 2e) and sharp (Figure 2f) kernels.

Figure 2: Auto-segmentation and observer augmentation.



The intersection between both erosion and image quality groups resulted in three features with excellent reproducibility, as follows: `gldm_GrayLevelNonUniformity`, `firstorder_Energy`, and `firstorder_TotalEnergy`.

The rf_mean model with these selected features performed with AUC = 1.0 in the training cohort and AUC = 0.84 (0.73-0.95) in the validation cohort, both at the subject level. Sensitivity and specificity in the validation cohort were 70.4% (51.5-84.2%) and 89.3 (72.3-96.3%), respectively (Table 3).

3.3 Physicians

In the physician's assessment, pneumologist performed with sensitivity and specificity of 81.5% (63.3-91.8%) and 67.9% (49.3-82.1%), respectively; while radiologist performed with sensitivity and specificity of 85.2% (67.5-94.1%) and 60.7% (42.4-76.43%), respectively. When physicians were assisted with both artificial intelligence models, their performance were sensitivity and specificity of, respectively, 85.2% (67.5-94.1%) and 85.7 (68.5-94.3) for pneumologist, and 88.9% (71.9-96.2%) and 96.4% (82.99.4%) for radiologist. The Cohen's Kappa coefficient between both physicians was $k = 0.66$ (0.46-0.86) without the support of machine learning and improved to 0.78 (0.61-0.94) with the use of machine learning. Therefore, the readings assisted by artificial intelligence improved the lower limit from moderate to substantial. Figure 3 depicts the overall readers' performance based on the AUC, and Figure 4 depicts the readers' reliability.

Even though we only found a significant difference between specificities for the radiologist reader with and without AI ($p < 0.05$), our results suggest that standardized and semi-automatic measures based on questionnaires and radiomics may assist physicians on the classification task.

Table 3 – Model performance on the validation dataset

Model	Sensitivity (95%CI)	Specificity (95%CI)	AUC (95%CI)
<i>Structured Report</i>	77.3 (56.7-89.9)	69.7 (52.7-82.6)	0.82 (0.69-0.93)
<i>Radiomic</i>	70.4 (51.5-84.2)	89.3 (72.3-96.3)	0.84 (0.73-0.95)

<i>Pneumologist</i>	81.5 (63.3-91.8)	67.9 (49.3-82.1)	0.84 (0.72-0.95)
<i>Radiologist</i>	85.2 (67.5-94.1)	60.7 (42.4-76.43)	0.78 (0.66-0.90)
<i>Pneumologist with AI</i>	85.2 (67.5-94.1)	85.7 (68.5-94.3)	0.92 (0.84-0.99)
<i>Radiologist with AI</i>	88.9 (71.9-96.2)	96.4 (82.3-99.4)	0.97 (0.92-1.00)

Fig. 3: Physicians and artificial intelligence performance.

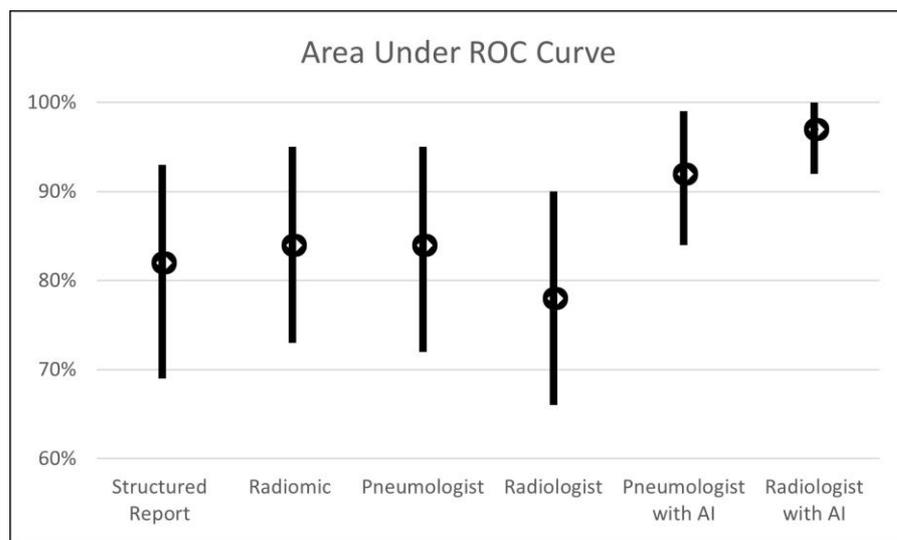
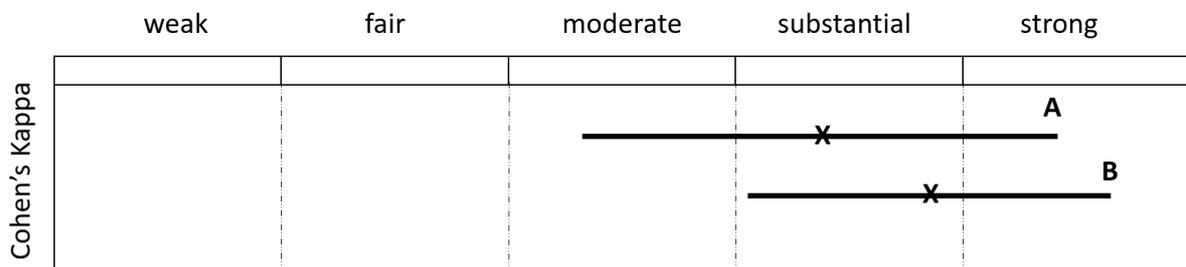


Fig. 4: Reader's reliability.



Physician's agreement (A) without and (B) with artificial intelligence.

DISCUSSION

This study shows that structured reports, radiomics, and machine learning algorithms provide a reliable classification of CT scans to identify COVID-19 individuals and better assist physicians in their decision. Our study characterized the CT pneumonia examinations using radiomics and a typically structured report questionnaire for lung CT pneumonia and AI models. Furthermore, we conducted a careful methodological and statistical design to avoid the pitfalls in the modeling workflow and analysis, thus promoting the results' reliability and repeatability [21,22].

Syeda et al. [14] published a systematic review on the role of machine learning in tackling coronavirus disease. Among forty studies that evaluated the diagnosis of COVID-19, none of them used imaging features extracted from standardized structured reports. Although several authors reported the application of AI for diagnoses of COVID-19 using CT, their works were based on image processing techniques and/ or clinical data.

We could not demonstrate that the model's performance was statistically higher than that achieved by the physicians, but the observed trends are optimistic. More importantly, they improved their reportings when assisted by these models. Interestingly, the results of our radiomic model are similar to Guiot J et al. [12], which achieved a sensitivity of 69.5% and a specificity of 91.6%. Liu H et al. [13] developed both clinical-radiological and clinical-radiomic models. They found sensitivity and specificity of 63% and 84% for their clinical-radiological model, which included the features of age, gender, neutrophil ratio, lymphocyte count, location, distribution, reticulation, and lung involvement. Its overall performance AUC = 0.83 (0.75–0.90) is also equivalent to our structured report model. Their clinical-radiomic model achieved AUC = 0.93 (0.85-1.00) with sensitivity and specificity of 85% and 90%, respectively. Although clinical data may provide additional predictive information, they are not available in many clinical settings, while our models depend only on simple scoring and automated image quantification.

Our work also confirmed the higher prevalence of typical features previously reported in COVID-19 cases (A1, A2, D1, D2) [1,23], demonstrating these patterns' added contribution and dependence on artificial intelligence models. These models' performance motivates a more extensive clinical evaluation based on a larger cohort designed to evaluate the gain in accuracy and inter- intra rater reliability achieved by radiologists of different levels of experience with the assistance of our algorithm.

Our study has some limitations. First, the limited sample size imposes a relatively large confidence interval of our results. Second, our radiomic model could be generalized only with the feature selection and extraction parameters used in this work. Finally, our algorithm shows two scores (radiomic and structured reports) therefore limiting the decision contribution to physicians when both scores are ambiguous (n=18/55).

The management of COVID-19 patients is multifaceted. Decision-making technologies to handle coronavirus disease in real-time are valuable tools to avoid its spread and rationalize resources. The proposed approach intends to standardize better, automate, and identify infected patients or discard pneumonia that is unlikely COVID-19. Our efforts are aligned with global demand to make minable data and artificial intelligence feasible and translational to clinical practice, as radiomics, reporting standardization, and statistical tools are becoming more widespread [24].

CONCLUSIONS

Radiomics, structured reports, and machine learning algorithms enable good classification of CT scans to identify COVID-19 individuals. In addition, this approach assisted physicians in standardizing better the classification and reporting of COVID-19 in computed tomography.

REFERENCES

1. Kwee TC and Kwee RM. Chest CT in COVID-19: what the radiologist needs to know. *RadioGraphics*. 2020; 40(7): 1848-65.

2. da Rosa Mesquita R, Francelino Silva Junior LC, Santos Santana FM, et al. Clinical manifestations of COVID-19 in the general population: systematic review. *Wien Klin Wochenschr.* 2021; 133(7-8):377-82.
3. Barbosa PNVP, Bitencourt AGV, Miranda GD, et al. Chest CT accuracy in the diagnosis of SARS-CoV-2 infection: initial experience in a cancer center. *Radiol Bras.* 2020; 53(4): 211-215.
4. Yasaka K and Abe. Deep learning and artificial intelligence in radiology: Current applications and future directions. *PLoS medicine.* 2018; 15(11): e1002707.
5. Dahl FA, Rama T, Hurlen P, et al. Neural classification of Norwegian radiology reports: using NLP to detect findings in CT-scans of children. *BMC Med Inform Decis Mak.* 2021; 21(1): 1-8.
6. Yi C, Tang Y, Ouyang R, et al. The added value of an artificial intelligence system in assisting radiologists on indeterminate BI-RADS 0 mammograms. *Eur Radiol.* 2022; 32(3): 1528-37.
7. Chan KS, Chan YM, Tan AHM, et al. Clinical validation of an artificial intelligence-enabled wound imaging mobile application in diabetic foot ulcers. *Int Wound J.* 2022; 19(1): 114-24.
8. Barbosa F, Maciel LMZ, Vieira EM, et al. Radiological reports: a comparison between the transmission efficiency of information in free text and in structured reports. *Clinics.* 2010; 65(1): 15-21.
9. Reiner BI. The challenges, opportunities, and imperative of structured reporting in medical imaging. *J Digit Imaging.* 2009; 22(6): 562-8.
10. Lakhani P, Prater AB, Hutson RK, et al. Machine learning in radiology: applications beyond image interpretation. *J Am Coll Radiol.* 2018; 5(2): 350-59.
11. Nguyen DH and Patrick JD. Supervised machine learning and active learning in classification of radiology reports. *J Am Med Inform Assoc.* 2014; 21(5): 893-901.
12. Guiot J, Vaidyanathan A, Deprez L, et al. Development and validation of an automated radiomic CT signature for detecting COVID-19. *Diagnostics.* 2020; 11(1): 41.
13. Liu H, Ren H, Wu Z, et al. CT radiomics facilitates more accurate diagnosis of COVID-19 pneumonia: compared with CO-RADS. *Journal of Translational Medicine.* 2021; 19(1): 1-12.
14. Syeda HB, Syed M, Sexton KW, et al. Role of machine learning techniques to tackle the COVID-19 crisis: Systematic review. 2021; *JMIR med inform.* 2021; 9(1): e23811.
15. Fan D, Zhou T, Ji G, et al. Inf-Net: Automatic COVID-19 Lung Infection Segmentation From CT Images. *IEEE Transactions on Medical Imaging.* 2020; 39(8): 2626-37.
16. Van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 2017; 77(21): e104-7.
17. Demsar J, Curk T, Erjavec A, et al. [Orange: Data Mining Toolbox in Python](#), *Journal of Machine Learning Research.* 2013; 14(35):2349–53.
18. Dean AG, Sullivan KM, Soe MM. OpenEpi: Open Source Epidemiologic Statistics for Public Health [www.OpenEpi.com] [updated 2013 Apr 06; cited in 2021 Aug 21].
19. Kundu S, Kers JG, Janssens ACJ. Constructing hypothetical risk data from the area under the ROC curve: modelling distributions of polygenic risk. *PLoS One.* 2016; 11(3): e0152359.
20. Sim J and Wright CC. The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Phys Ther.* 2005; 85(3): 257–68.
21. Vanwinckelen G and Blockeel H. On estimating model accuracy with repeated cross-validation. *Benelearn 2012: Proceedings of the 21st Belgian-Dutch conference on machine learning.* 2012; 39-44.
22. Ying X. An overview of overfitting and its solutions. *J. Phys.: IOP Conf. Ser.* 2019; 1168(2): 022022.

23. Adams HJ, Kwee TC, Yakar D, et al. Chest CT imaging signature of coronavirus disease 2019 infection: in pursuit of the scientific evidence. *Chest*. 2020; 158(5): 1885-95.
24. Buvat I and Orhac F. The TRUE Checklist for Identifying Impactful Artificial Intelligence–Based Findings in Nuclear Medicine: Is It True? Is It Reproducible? Is It Useful? Is It Explainable? *J Nuc Med*. 2021; 62(6): 752-54.

6. DISCUSSÃO

Neste projeto desenvolvemos e implementamos tecnologias da informação e métodos computacionais para otimização dos processos na saúde. A incorporação das novas metodologias no software WBA permitiu validar os componentes tecnológicos em ambiente operacional, visando o atingimento de maturidade tecnológica TRL7, que é quando uma tecnologia é demonstrada em um ambiente clínico (norma ISO 16290:2013). A figura 5 orienta sobre os resultados em cada etapa do trabalho.

A implementação do sistema de gerenciamento de filas demonstrou, na percepção dos profissionais usuários do *software*, que 86,7% dos usuários consideraram que a solução de gerenciamento de filas da plataforma agregou mais segurança aos pacientes, enquanto 13,3% dos usuários consideraram indiferente. Em relação aos ganhos de produtividade, 83,3% dos usuários relataram ganhos no atendimento dos pacientes, enquanto 16,7% dos usuários julgaram indiferente (anexo A, anexo C) (PONCE *et al.*, 2020). Estes resultados demonstram o potencial de melhorias que podem ser alcançadas na adoção de ferramentas aderentes às especificidades dos processos de saúde, permitindo mais flexibilidade e mais eficiência na gestão de processos (ZANG *et al.*, 2009). Para o segmento de diagnóstico por imagens, faz-se necessário integrar os softwares de gestão com uma solução de PACS para otimizar os ganhos de eficiência (AYAL and SEIDMAN *et al.*, 2009). Durante nosso estudo, evoluímos a ferramenta PACS para o conceito de PACS de Larga Escala, viabilizando uma infraestrutura tecnológica capaz de gerenciar mais de 20.000 exames por ano, validado para as modalidades SPECT, PET, CT, densitometria óssea, raios-X e ultrassom (MANSOORI *et al.*, 2012).

As ferramentas de aceleração de exames buscam otimizar a produtividade dos serviços de saúde e prover melhor satisfação e conforto aos pacientes, além de reduzir custos operacionais com a otimização da capacidade instalada. Desenvolvemos um algoritmo que permite particionar imagens convencionais de SPECT em novas imagens que simulam terem sido realizadas em menor tempo (artigo 1 e anexo C) (QUEIROZ *et al.*, 2022). Este *software* permite não apenas facilitar a avaliação de novas tecnologias de aceleração de imagens (ao simular imagens com menor tempo, sem necessidade de expor o paciente a uma nova aquisição de imagens), como ainda facilitará o desenvolvimento de novos modelos com *deep-learning* (dada a alta capacidade de simular imagens de variadas qualidades), reduzindo também a dependência de tecnologia importada (VALENTA *et al.*, 2010; LECCHI *et al.*, 2017; OLIA *et al.*, 2022).

Nas aplicações computacionais para acesso a biomarcadores e apoio à decisão, faz-se necessário entender e estudar as metodologias de padronização e seleção (escolha) dos atributos de imagens conforme sua robustez. Uma das formas de análise, geralmente complementar, refere-se a estudos de repetitividade com objetos simuladores (*phantoms*). Em nosso estudo, caracterizamos a repetitividade de atributos radiômicos usando o simulador IEC/NEMA (artigo 2) (MACHADO *et al*, 2021) e estudamos métodos de harmonização e robustez (OLIVEIRA *et al*, 2022). Com isso, avaliamos correspondências na literatura sobre a consistência (reprodutibilidade, harmonização, correlação clínica) dos atributos radiômicos encontrados e aplicamos um estudo de caso em imagens de 18F-FDG-PET/CT de pacientes com linfoma Hodgkin (artigo 4) (MACHADO *et al*, 2022) e pacientes com COVID-19 (artigo 6 e anexo C) (MACHADO *et al*, 2022). Em nosso experimento com *phantom* para verificar a variabilidade dos atributos radiômicos PET, utilizamos um modelo simples, com esferas homogêneas, cujo resultado não é generalizável para condições em que o tumor se apresenta com captação mais heterogênea do radiofármaco, conforme Pfaehler e colegas (2019) demonstraram em seu estudo com *phantom* criado a partir de uma impressora 3D para simular casos de câncer de pulmão. Entretanto, nosso experimento contribui para demonstrar o comportamento da variabilidade dos atributos radiômicos em função de outros métodos de discretização dos valores de pixel.

Ressalta-se a necessidade de se estabelecer um workflow para segmentação de imagens, extração e seleção dos atributos radiômicos (Fig. 3). A escolha do conjunto de ferramentas necessárias dependerá da preferência da equipe de pesquisa e dos recursos disponíveis. Neste projeto usamos os *softwares* ImageJ, LIFEx e PyRadiomics, estabelecendo diferentes workflows para extração de características radiômicas (artigos 3-6) (¹MACHADO *et al*, 2022; OLIVEIRA, *et al*, 2022; ²MACHADO *et al*, 2022; ALENCAR *et al*, 2022). Apesar de todas estas opções apresentarem-se úteis, a incorporação de métodos baseados em linguagem Python possui a vantagem de facilitar a interoperabilidade com outros *softwares* (VANROSSUM, G *et al*, 1995). Por fim, optou-se por adotar a biblioteca PyRadiomics, que possui linguagem baseada em Python (GRIETHUYSEN, *et al*, 2017). Com isso, desenvolvemos aplicações de fronteira com *deep-learning* para segmentação de imagens (artigo 3, artigo 6 e anexo B) (OLIVEIRA, *et al*, 2022; MACHADO *et al*, 2022), métodos para obtenção dos biomarcadores e métodos *machine-learning* para a criação de modelos preditivos para apoio diagnóstico de linfoma Hodgkin (artigo 4) (MACHADO *et al*, 2022), colo de útero (artigo 5) (ALENCAR *et al*, 2022) e COVID-19 (artigo 6 e anexo C) (MACHADO *et al*, 2022). O desenvolvimento de algoritmos com redes neurais (*deep-learning*) para segmentação automática de imagens (artigo

3, artigo 6, anexo B e anexo C) é um importante recurso desenvolvido neste trabalho, pois automatiza o processo manual e elimina o viés do observador (operador-dependente), pois muitos biomarcadores são sensíveis à região delimitada da imagem (ex.: geometria da segmentação, região considerada, fatores de ponderação) (PFAEHLER *et al*, 2019).

Embora este conjunto de tecnologias tenha o potencial de apoiar a decisão médica, possibilitando a melhoria da qualidade de vida das pessoas, redução da hospitalização, da recidiva de doença e da mortalidade; cabe salientar a necessária validação clínica das tecnologias propostas, delimitando suas fragilidades, oportunidades e níveis de evidência. No modelo preditivo para apoio ao diagnóstico de linfoma Hodgkin (artigo 4) (MACHADO *et al*, 2022), devido ao tamanho amostral limitado, não foi possível realizar a seleção de biomarcadores e a validação do modelo em conjuntos de pacientes independentes, conforme preconiza as melhores práticas na área (BUVAT & ORLHAC, 2021), possivelmente resultado em viés de informação. Ambos os estudos apresentados nos artigos 4 e 5 foram realizados em um único equipamento de aquisição de imagens, necessitando de validação externa. Por outro lado, no artigo 6 tanto a seleção de biomarcadores quanto o treinamento dos modelos IA foram realizados em um grupo de pacientes independente do grupo de validação do modelo IA. Adicionalmente, foram incluídos três centros diagnósticos com três modelos de equipamentos distintos, incrementando a qualidade dos resultados. É importante destacar ainda que, além da melhora na acurácia diagnóstica observada no artigo 6, a IA também possibilitou uma melhora na concordância entre observadores. Todavia, com a constante modificação das características epidemiológicas da doença, faz-se necessário avaliar a tecnologia em novos pacientes após a consolidação da vacinação no Brasil.

Nossos resultados podem impactar muitos pacientes, como aqueles encaminhados para realizar CT de tórax com suspeita de COVID-19, que apesar da redução de casos decorrente da vacinação, continua afetando milhares de indivíduos por dia; pacientes oncológicos que realizam exames de PET/CT, cujos médicos podem acessar informações de textura tumoral, que está associada a um estágio mais avançado da doença (necrose); pacientes que realizam exames de medicina nuclear, que poderão usufruir de exames mais rápidos e/ou com menor dose de radiação; e outros pacientes encaminhados para procedimentos em departamentos de diagnóstico por imagens, que se beneficiarão de tecnologias integradas que melhoram a eficiência operacional dos serviços de saúde, resultando em mais segurança e mais produtividade.

7. CONCLUSÃO

Neste trabalho foi desenvolvido e implementado um conjunto de tecnologias da informação e métodos computacionais para melhorar a gestão e a assistência aos pacientes em serviços de diagnóstico por imagens:

a) A produtividade de serviços de saúde aumentou com a disponibilização de um software especialista com conceito WBA.

b) Exames podem ser realizados em menor tempo e/ou com menor dose de radiação utilizando o método para simulação de imagens desenvolvido e validado para protocolos de exames SPECT.

c) Foi desenvolvido e validado um workflow para estudos com radiômica, contemplando as etapas de segmentação automática de imagens, extração, padronização e seleção de biomarcadores a serem usadas na descoberta de novas associações (*Knowledge Discovery – KDD*).

d) Foram demonstrados novos biomarcadores para apoio à decisão em exames de 18F-FDG-PET/CT oncológico (linfoma de Hodgkin e câncer de colo de útero) e CT de tórax (Covid-19).

e) Um novo *software* desenvolvido para apoio ao diagnóstico da COVID-19 utilizando imagens de CT, melhorou a acurácia e concordância entre observadores na suspeita diagnóstica de pneumonia por COVID-19.

8. PERSPECTIVAS

As ferramentas desenvolvidas em uma plataforma de gestão no conceito WBA integrada a um PACS de Larga Escala é facilmente escalável para milhares de serviços de saúde e isto pode viabilizar a implementação de novos escores para a prática da medicina de precisão. Os resultados deste trabalho abrem perspectivas para a implementação do *software* de gestão WBA para outras especialidades médicas. O método de gerenciamento de filas validado para as especialidades de radiologia e medicina nuclear será abordado futuramente nas áreas de cardiologia e oncologia. Ainda na área de imagens, serão implementadas novas funcionalidades no *software* para contemplar novos fluxos de trabalho que inclua o médico assistente e outros profissionais externos aos serviços de saúde. Por exemplo, os resultados de exames diagnósticos usualmente disponibilizados pela internet para o paciente, serão disponibilizados também para o médico assistente (incluindo as ferramentas de IA), para engajar o médico assistente nesta etapa do processo de decisão, ampliando o alcance das soluções de escores de risco com radiômica e IA. Com a eminente implementação de políticas de proteção do sigilo do paciente reforçado pela Lei Geral de Proteção de Dados (LGPD), serão trabalhadas etapas de validação de identidade no usuário quando for necessária a identificação do paciente.

No campo da otimização do tempo de aquisição de imagens, o método desenvolvido para simulação de imagens com ruído será usado futuramente para treinar redes neurais de aprendizado profundo (*deep-learning*), visando uma redução do tempo dos exames SPECT, PET e RM. Esta iniciativa poderá acarretar numa mudança de paradigma para a realização destes exames, reduzindo o tempo do paciente nos serviços de saúde, aumentando o seu conforto e permitindo que os serviços de saúde alcancem um novo patamar na produtividade e eficiência operacional.

Com o workflow desenvolvido para extração de biomarcadores radiômicos, serão implementadas análises em outras doenças visando tanto validade das descobertas como incluir a informação radiômica em um repositório de variáveis clínicas (biobanco digital) para análises epidemiológicas, permitindo a descoberta de novas associações entre fatores de risco (ex.: diabetes, histórico familiar, exames, radiômica) e desfechos clínicos (ex.: hospitalização, recidiva de doença, morte). Esta iniciativa está alinhada ao campo de estudos do *BigData*, possibilitando analisar grandes massas de dados a partir da agregação de repositórios de dados. Com isto, as tecnologias desenvolvidas alcançarão maior utilidade, por exemplo apoiando pesquisadores e gestores da saúde na implementação de políticas e novos protocolos de saúde.

O auxílio de algoritmos com inteligência artificial para apoio à decisão, associando-se fatores clínicos e técnicas de radiômica, oferecem informações que são complementares na explicação da doença, enquanto a IA fornecerá o poder de combinar estas informações com maior acurácia preditiva, fornecendo uma ferramenta sem precedentes na prática da medicina personalizada.

Por meio do projeto FINEP Startups IA (Edital 04/2022), métodos IA com processamento de linguagem natural (NLP) serão usados para extrair o conhecimento armazenado em arquivos de texto com linguagem natural pelos serviços de saúde. Este método poderá auxiliar na coleta de grandes volumes de informações para identificar associações, facilitar a coleta de dados presentes em sistemas obsoletos, e auxiliar no desenvolvimento de modelos multivariados. Biomarcadores extraídos por diferentes métodos, por exemplo, a fração de ejeção extraída da cintilografia do miocárdio e a fração de ejeção extraída pelo exame de ecocardiograma também podem ser usados em técnicas de harmonização (artigo 3), a fim de comparar medidas numa mesma escala. Ainda por meio desta iniciativa, serão contemplados no sistema PACS/Viewer novos biomarcadores de imagens como a composição corporal extraída da CT (Anexo B) e a textura tumoral de imagens PET (Artigos 4 e 5). Assim, a arquitetura tecnológica demonstrada para a COVID-19 (figura 4) será expandida para acessar biomarcadores e assinaturas IA de outros exames PET e CT.

Será também explorada a correlação de biomarcadores radiômica com biomarcadores moleculares para substituição de procedimentos mais invasivos. Existe ainda um amplo espaço para aplicações de radiômica em outras modalidades de imagens médicas, como ultrassom, patologia, ressonância magnética, fotografias de feridas de úlceras falciformes. A rápida viabilidade de aplicações de *softwares* e inteligência médica na saúde poderá desencadear oportunidades de geração de novas *startups* e inovações na saúde.

REFERÊNCIAS

ACR accreditation. Disponível em: <www.acraccreditation.org>. Acesso em: 19/09/2022.

ALMEIDA, M.; PEREIRA, L.; Ren, T.; CAVALCANTI, C.; SIJBERS, J. The Gated Recurrent Conditional Generative Adversarial Network (GRC-GAN): application to denoising of low-dose CT images. **In: 2021 34th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). IEEE**, 2021. p. 129-135. ISSN 0278-0062

ARMSTRONG, I.; PARTHIBAN, A.; JAMES, J.; TONGE, C.; LAWSON, R. Reduced-count myocardial perfusion SPECT with resolution recovery. **Nuclear Medicine Communications**, v. 33, n. 2, p. 121-129, 2012. ISSN 1473-5628.

ASH, J. and BATES, D. Factors and forces affecting EHR system adoption: report of a 2004 ACMI discussion. **Journal of the American Medical Informatics Association**, v. 12, n. 1, p. 8-12, 2005. ISSN 1067-5027.

AYAL, M. and SEIDMAN, A. An empirical investigation of the value of integrating enterprise information systems: the case of medical imaging informatics. **Journal of Management Information Systems**, v. 26, n. 2, p. 43-68, 2009. ISSN 0742-1222.

BARBOSA PNVP, BITENCOURT AGV, MIRANDA GD, et al. Chest CT accuracy in the diagnosis of SARS-CoV-2 infection: initial experience in a cancer center. **Radiologia Brasileira**, v. 53, p. 211-215, 2020. ISSN 0100-3984

BLAZONA, B and KONCAR, M. HL7 and DICOM based integration of radiology departments with healthcare enterprise information systems. **International Journal of Medical Informatics**, v. 76, n.3, p. 425-432, 2007. ISSN 1386-5056.

BRAY F, FERLAY J, SOERJOMATARAM I, SIEGEL R, TORRE L, JEMAL A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA Cancer J Clin**. (2018) 68: 394–424. doi: 10.3322/caac.21492. ISSN 1542-4863

Buvat I & Orhac F. The TRUE Checklist for Identifying Impactful Artificial Intelligence–Based Findings in Nuclear Medicine: Is It True? Is It Reproducible? Is It Useful? Is It Explainable? **J Nuc Med**. 2021; 62(6): 752-54. ISSN 2159-662X

CBOK, BPM. Guide to the business process management common body of knowledge. Versão, v. 2, p. 2009, 2009.

COOK, G.; SIDDIQUE, M.; TAYLOR, B.; YIP, C.; CHICKLORE, S.; GOH, V. Radiomics in PET: principles and applications. **Clinical and Translational Imaging**, v. 2, n. 3, p. 269-276, 2014. ISSN 2281-7565

COTTEREAU AS, VERSARI A, LOFT A, et al. Prognostic value of baseline metabolic tumor volume in early-stage Hodgkin lymphoma in the standard arm of the H10 trial. **Blood**. 2018;131(13):1456-1463. ISSN: 0006-4971

CRANDALL, J.; FRAUM, T.; LEE, M.; JIANG, L.; GRIGSBY, P.; WAHL, R. Repeatability of 18F-FDG PET radiomic features in cervical cancer. **Journal of Nuclear Medicine**, v. 62, n. 5, p. 707-715, 2021. ISSN 2159-662X

DA CUNHA JÚNIOR, Ademar Dantas et al. Adipose tissue radiodensity: A new prognostic biomarker in people with multiple myeloma. **Nutrition**, v. 86, p. 111141, 2021. ISSN 0899-9007

DESHMANE, A.; GULANI, V.; GRISWOLD, M.; SEIBERLICH, N. Parallel MR imaging. **Journal of Magnetic Resonance Imaging**, v. 36, n. 1, p. 55-72, 2012. ISSN 1522-2586.

EBADI, Maryam et al. Visceral adipose tissue radiodensity is linked to prognosis in hepatocellular carcinoma patients treated with selective internal radiation therapy. **Cancers**, v. 12, n. 2, p. 356, 2020. ISSN 2072-6694

FANNY O.; BOUGHADAD, S.; PHILIPPE, C.; *et al.* A postreconstruction harmonization method for multicenter radiomic studies in PET. **Journal of Nuclear Medicine**, v. 59, n. 8, p. 1321-1328, 2018. ISSN 2159-662X

FERLAY J, SOERJOMATARAM I, DIKSHIT R, ESER S, MATHERS C, REBELO M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. **International journal of cancer**, v. 136, n. 5, p. E359-E386, 2015. ISSN 1097-0215

FIOCRUZ. Estudo analisa registro de óbitos por Covid-19 em 2020. Disponível em <<https://portal.fiocruz.br/noticia/estudo-analisa-registro-de-obitos-por-covid-19-em-2020>> Acessado em 19.11.2022

HOEKSTRA, C. J.; PRUIM, J.; WILLEMSSEN, A.; ARENDS, B.; KOTZERKE, J.; BOCKISCH, A.; BEYER, T.; CHITI, A.; KRAUSE, B. J. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. **European journal of nuclear medicine and molecular imaging**, p. 1-26, 2014. ISSN 1619-7070.

HUANG, Y.; WEI, L.; HU, Y.; SHAO, N.; LIN, Y.; HE, S.; SHI, H.; ZHANG, X.; LIN, Y. Multi-parametric MRI-based radiomics models for predicting molecular subtype and androgen receptor expression in breast cancer. **Frontiers in Oncology**, v. 11, 2021. ISSN 2234-943X

IM HJ, BRADSHAW T, SOLAIYAPPAN M, CHO SY. Current methods to define metabolic tumor volume in positron emission tomography: which one is better?. **Nuclear medicine and molecular imaging**, v. 52, n. 1, p. 5-15, 2018. ISSN 1619-7089

IROJU, O.; SORIYAN, A.; GAMBO, I; OLALEKE, J. Interoperability in healthcare: benefits, challenges and resolutions. **International Journal of Innovation and Applied Studies**, v. 3, n. 1, p. 262-270, 2013. ISSN 2028-9324

ISO/FDIS 16290:2013(E) Space systems - Definition of the Technology Readiness Levels (TRLs) and their criteria of assessment. **International Organization for Standardization**, Switzerland, 2013. 12p.

JAMBOR, I. Optimization of prostate MRI acquisition and post-processing protocol: a pictorial review with access to acquisition protocols. **Acta Radiologica Open**, v. 6, n. 12, p. 1-14, 2017. ISSN 2058-4601.

KAALEP, A.; SERA, T.; OYEN, W.; KRAUSE, B.; CHITI, A.; LIU, Y.; BOELLAARD, R. EANM/EARL FDG-PET/CT accreditation-summary results from the first 200 accredited imaging systems. **European journal of nuclear medicine and molecular imaging**, v. 45, n. 3, p. 412-422, 2018. ISSN 1619-7070.

KANOUN, Salim et al. Influence of software tool and methodological aspects of total metabolic tumor volume calculation on baseline [18F] FDG PET to predict survival in Hodgkin lymphoma. **PloS one**, v. 10, n. 10, p. e0140830, 2015. ISSN 1932-6203.

KAPLAN, S. and ZHU, Y. Full-dose PET image estimation from low-dose PET image using deep learning: a pilot study. **Journal of digital imaging**, v. 32, n. 5, p. 773-778, 2019. ISSN 1618-727X

KHIEWVAN B, TORIGIAN DA, EMAMZADEHFARD S, PAYDARY K, SALAVATI A, HOUSHMAND S, et al. Update of the role of PET/CT and PET/MRI in the management of patients with cervical cancer. **Hellenic journal of nuclear medicine**, v. 19, n. 3, p. 254-268, 2016. ISSN 1790-5427

KOENIG, Inke R. et al. What is precision medicine? **European respiratory journal**, v. 50, n. 4, 2017.

KOPEMAN, D.; van OSCH, J. A.; JAGER, P. L.; TENBERGEN, C. J.; KNOLLEMA, S.; SLUMP, C. H.; van DALEN, J. A. Technical note: how to determine the FDG activity for tumor PET imaging that satisfies European guidelines. **European journal of nuclear medicine and molecular imaging physics**, v. 3, n. 1, p. 22, 2016. ISSN 2197-7364 (Online).

KWEE TC and KWEE RM. Chest CT in COVID-19: what the radiologist needs to know. **Radiographics**, v. 40, n. 7, p. 1848, 2020. ISSN 1527-1323

LECCHI, M.; MARTINELLI, I.; ZOCCARATO, O.; MAIOLI, C.; LUCIGNANI, G.; DEL SOLE, A. Comparative analysis of full-time, half-time, and quarter-time myocardial ECG-gated SPECT quantification in normal-weight and overweight patients. **Journal of Nuclear Cardiology**, v. 24, n. 3, p. 876-887, 2017. ISSN 1532-6551.

MACHADO, M.; MENEZES, V.; NAMÍAS, M.; VIEIRA, N.; QUEIROZ, C.; MATHEOD, R.; ALESSIO, A.; OLIVEIRA, M. Protocols for harmonized quantification and noise reduction in low-dose oncologic 18F-FDG PET/CT imaging. **Journal of Nuclear Medicine Technology**, v. 47, n. 1, p. 47-54, 2019. ISSN 0091-4916.

MACHADO, M. A. D.; MENEZES, V. O.; NAMIAS, M; VIEIRA, N. S.; QUEIROZ, C. C.; OLIVEIRA, M. L. Technology Innovation Based on Specific Nuclear Medicine Software for PET/CT Protocol Standardization and Quality Assurance Management. **International Symposium – Quantification in Medical and Preclinical Imaging: state of the art and future developments**. Groningen, Holanda, 2016.

MACHADO, M.; PITA, J.; NETTO, E. M. PET Radiomic Features Variability: A Phantom Study on the Influence of Reconstruction and Discretization Method. **Revista Brasileira de Física Médica**. v.15, 2021. ISSN 1984-9001.

MACHADO, M. A. D., QUEIROZ, C. C, PITA, J. L. P, et al. Prediction of Treatment Failure in Hodgkin Lymphoma: A Machine Learning Approach in Baseline 18F-FDG PET/CT. **Revista Brasileira de Física Médica**. Submetido em outubro/2022.

MACHADO, M. **Metodologia para otimização de protocolo PET/CT harmonizado de baixa dose em tumores sólidos com reconstrução Point Spread Function**. Orientador: Mércia Liane. 2017. 79f. Dissertação (Mestrado). Programa de Pós-Graduação em Tecnologias Energéticas e Nucleares. Centro Regional de Ciências Nucleares do Nordeste, Departamento de Energia Nuclear, Universidade Federal de Pernambuco. Disponível em <<https://repositorio.ufpe.br/handle/123456789/25171>>

MANSOORI, B.; ERHARD, K.; SUNSHINE, J. Picture Archiving and Communication System (PACS) implementation, integration & benefits in an integrated health system. **Academic radiology**, v. 19, n. 2, p. 229-235, 2012. ISSN 1076-6332.

MARTIN, C. and SOOKPENG, S. Setting up computed tomography automatic tube current modulation systems. **Journal of Radiological Protection**, v. 36, n. 3, p. R74, 2016. ISSN 0952-4746.

MILGROM SA, ELHALAWANI H, LEE J, et al. A PET Radiomics Model to Predict Refractory Mediastinal Hodgkin Lymphoma. **Scientific reports**, v. 9, n. 1, p. 1-8, 2019. ISSN: 2045-2322

MENEZES, V.; MACHADO, M.; QUEIROZ, C.; SOUZA, O. S.; d'ERRICO, F.; NAMIAS, M.; LAROCCA, T. F.; SOARES, M. B. P. Optimization of oncological 18F-FDG PET/CT imaging based on a multiparameter analysis. **Medical physics**, v. 43, n. 2, p. 930-938, 2016. ISSN 0094-2405.

MENEZES, V.; MACHADO, M.; QUEIROZ, C.; REINANDE, I.; VIEIRA, L.; GAMA, J. How does Lean Six Sigma method improve healthcare practice in nuclear medicine departments? A successful case of dedicated software applications in oncological PET/CT. **Journal of Nuclear Medicine** v.59 (supplement 1) 1013, 2018. ISSN 0161-5505

METTLER J, MÜLLER H, VOLTIN CA, BAUES C, KLAESER B, MOCCIA A, BORCHMANN P, ENGERT A, KUHNERT G, DRZEZGA AE, DIETLEIN M, KOBE C. Metabolic Tumour Volume for Response Prediction in Advanced-Stage Hodgkin Lymphoma. **J Nucl Med**. 2018 Jun 7;60(2):207–11. ISSN: 0161-5505

NAMÍAS, M.; BRADSHAW, T.; MENEZES, V.; MACHADO, M.; JERAJ, R. A novel approach for quantitative harmonization in PET. **Physics in Medicine & Biology**, v. 63, n. 9, p. 095019, 2018. ISSN 1361-6560

NANCE, J.; MEENAN, C.; NAGY, P. The future of the radiology information system. **American Journal of Roentgenology**, v. 200, n. 5, p. 1064-1070, 2013. ISSN 1546-3141.

NIOCHE, Christophe et al. LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. **Cancer research**, v. 78, n. 16, p. 4786-4789, 2018. ISSN 1538-7445.

OLIA, N.; KAMALI-ASL, A.; TABRIZI, S.; GERAMIFAR, P.; SHEIKHZADEH, P.; FARZANEFAR, S.; ARABI, H.; ZAIDI, H. Deep learning-based denoising of low-dose SPECT myocardial perfusion images: quantitative assessment and clinical performance. **European journal of nuclear medicine and molecular imaging**, v. 49, n. 5, p. 1508-1522, 2022. ISSN 1619-7070

PADHANI, A. and OLLIVIER, L. The RECIST criteria: implications for diagnostic radiologists. *The British journal of radiology*, v. 74, n. 887, p. 983-986, 2001. ISSN 2053-1303

PARK, J.; KIM, H.; KIM, D.; PARK, S.; KIM, J.; CHO, S.; KIM, J. A systematic review reporting quality of radiomics research in neuro-oncology: toward clinical utility and quality improvement using high-dimensional imaging features. **BMC cancer**, v. 20, n. 1, p. 1-11, 2020. ISSN 1471-2407

PFAEHLER, Elisabeth et al. Repeatability of 18F-FDG PET radiomic features: A phantom study to explore sensitivity to image reconstruction settings, noise, and delineation method. **Medical physics**, v. 46, n. 2, p. 665-678, 2019. ISSN 2473-4209

PFAEHLER, E.; MESOTTEN, L.; ZHOVANNIK, I.; PIEPLENBOSCH, S.; THOMEER, M.; VANHOVE, K.; ADRIAENSENS, P.; BOELLAARD, R. Plausibility and redundancy analysis to select FDG-PET textural features in non-small cell lung cancer. **Medical physics**, v. 48, n. 3, p. 1226-1238, 2021. ISSN 2473-4209

PONCE, D.; MACHADO, M.; MENEZES, M.; QUEIROZ, Q.; ALCANTARA, A.; SENA, T.; MARTINIANO, M.; VIGÁRIO, V. Plataforma para Gerenciamento de Filas e Rastreabilidade de Processos para Aumentar a Produtividade e a Segurança do Paciente. **Anais do XVII Congresso Brasileiro de Informática em Saúde**, p. 54-55, 2020. ISSN-2178-2857.

QUEIROZ, C.; MACHADO, M.; XIMENES, A.; PINO, A.; NETTO, E. Technical note: Partitioning of gated single photon emission computed tomography raw data for protocols optimization. **J Appl Clin Med Phys**, 23:e13508. 2022. ISSN 1526-9914.

SACHS, Peter B. et al. CT and MR protocol standardization across a large health system: providing a consistent radiologist, patient, and referring provider experience. **Journal of digital imaging**, v. 30, n. 1, p. 11-16, 2017. ISSN 1718-727X

SANAAT, A.; SHIRI, I.; ARABI, H.; MAINTA, I.; NKOULOU, R.; ZAIDI, H. Deep learning-assisted ultra-fast/low-dose whole-body PET/CT imaging. **European journal of nuclear medicine and molecular imaging**, v. 48, n. 8, p. 2405-2415, 2021. ISSN 1619-7070

SYEDA HB, SYED M, SEXTON KW, et al. Role of machine learning techniques to tackle the COVID-19 crisis: systematic review. **JMIR medical informatics**, v. 9, n. 1, p. e23811, 2021. ISSN 2291-9694

STRIMBU, Kyle; TAVEL, Jorge A. What are biomarkers?. **Current Opinion in HIV and AIDS**, v. 5, n. 6, p. 463, 2010. ISSN 1746-6318

SUI, H.; LIU, L.; LI, X.; ZUO, P.; CUI, J.; MO, Z. CT-based radiomics features analysis for predicting the risk of anterior mediastinal lesions. **Journal of Thoracic Disease**, v. 11, n. 5, p. 1809, 2019. ISSN 2072-1439

SUNG H, FERLAY J, SIEGEL RL, LAVERSANNE M, SOERJOMATARAM I, JEMAL A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: a cancer journal for clinicians**, v. 71, n. 3, p. 209-249, 2021. ISSN 1542-4863

TIXIER, F.; Le REST, C.; HATT, M.; ALBARGHACH, N.; PRADIER, O.; METGES, J.; CORCOS, L.; VISVIKIS, D. et al. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. **Journal of Nuclear Medicine**, v. 52, n. 3, p. 369-378, 2011. ISSN 0161-5505

VALENTA, I.; TREYER, V.; HUSMANN, L.; GAEMPERLI, O.; SCHINDLER, M.; HERZOG, B.; VEIT-HEIBACH, P.; BUECHEL, R.; NKOULOU, R.; PAZHENKOTTIL, A.; KAUFMANN, P. New reconstruction algorithm allows shortened acquisition time for myocardial perfusion SPECT. **European journal of nuclear medicine and molecular imaging**, v. 37, n. 4, p. 750-757, 2010. ISSN 1619-70789.

VALLIÈRES, M.; ZWANENBURG, A.; BADIC, B.; Le REST, C.; VISVIKIS, D.; HATT, M. Responsible radiomics research for faster clinical translation. **Journal of Nuclear Medicine**, v. 59, n. 2, p. 189-193, 2018. ISSN 0161-5505

VAN GRIETHUYSEN, Joost JM et al. Computational radiomics system to decode the radiographic phenotype. **Cancer research**, v. 77, n. 21, p. e104-e107, 2017. ISSN 1538-7445.

VANROSSUM, G. Python reference manual. Department of Computer Science [CS], n. R 9525, 1995.

WAHL, R. L.; JACENE, H.; KASAMON, Y.; LODGE, M. A. From RECIST to PERSIST: evolving considerations for PET response criteria in solid tumors. **Journal of Nuclear Medicine**. v. 50, n. suppl 1 , p. 122S-150S, 2009. ISSN 0161-5505.

YANG, G.; YU, S.; DONG, H.; SLABAUGH, G.; DRAGOTTI, P.; YE, X.; LIU, F.; ARRIDGE, S.; KEEGAN, J.; GUO, Y.; FIRMIN, D. DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. **IEEE transactions on medical imaging**, v. 37, n. 6, p. 1310-1321, 2017. ISSN 0278-0062

ZANG, J.; LU, X.; NIE, H; HUANG, Z.; van der Aalst, P.; Radiology information system: a workflow-based approach. **International Journal of Computer Assisted Radiology and Surgery**, v.4, p. 509-516, 2009. ISSN 1861-6429.

ZWANENBURG, A *et al.* The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. **Radiology**, v. 295, n. 2, p. 328-338, 2020. ISSN 1527-1315

ANEXOS

Anexo A – Apresentação em Congresso

Plataforma para Gerenciamento de Filas e Rastreabilidade de Processos para Aumentar a Produtividade e a Segurança do Paciente

Trabalho publicado nos anais do XVII Congresso Brasileiro de Informática em Saúde, dezembro de 2020.

Plataforma para Gerenciamento de Filas e Rastreabilidade de Processos para Aumentar a Produtividade e a Segurança do Paciente

Dayane Ponce¹, Marcos Machado^{1,2,3,4}, Vinicius Menezes^{1,4,5,6}, Cleiton Queiroz^{1,3,4,6}, Anna Alcântara¹, Thiago Sena¹, Anderson Martiniano¹, Adriano Vigário¹,

¹Nuclearis Sistemas em Medicina Nuclear, Radiologia e Oncologia, Salvador, BA

²Hospital das Clínicas da Universidade Federal da Bahia, Salvador, BA

³Hospital Cárdio Pulmonar, Salvador, BA

⁴Radtec Serviços em Física Médica, Salvador, BA

⁵Hospital das Clínicas da Universidade Federal de Pernambuco, Recife, PE

⁶Hospital São Rafael, Salvador, BA

⁷Hospital das Clínicas da Universidade Federal de Alagoas, Maceió, AL

RESUMO

Segurança e produtividade no atendimento ao paciente é uma preocupação constante no processo de gestão da saúde. Sendo assim, é necessário prover tecnologias digitais que contemplem as especificidades dos diferentes serviços de saúde, que seja capaz de monitorar o paciente durante toda a sua jornada no serviço de saúde; com objetivo de aumentar a produtividade e minimizar a ocorrência de eventos adversos e erros na rotina de atendimento. Neste trabalho apresentamos a experiência dos usuários com o sistema Nuclearis, uma plataforma aderente a processos clínicos e administrativos baseado num workflow sob medida.

PALAVRAS-CHAVE:

Produtividade; Rastreabilidade; Gestão pela Qualidade.

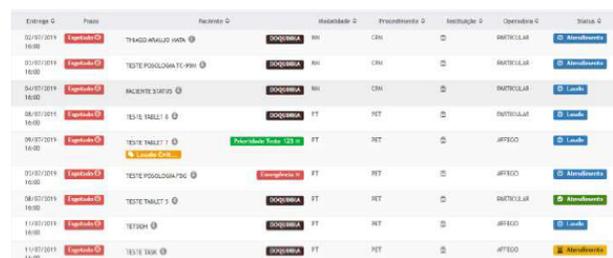
INTRODUÇÃO

O Programa Nacional de Segurança do Paciente (PNSP) foi criado para contribuir para a qualificação do cuidado em saúde no país¹. A magnitude dos eventos adversos e o pouco investimento necessário para a implantação dos protocolos de segurança do paciente justificam a implantação destes protocolos². A maioria dos eventos adversos estão relacionados a problemas operacionais ou a erros de medicações, o qual estima-se que a incidência em hospitais seja de 9,2%, sendo metade destes preveníveis³.

A plataforma Nuclearis tem demonstrado resultados satisfatórios no uso de inteligência de processos para apoio à tomada de decisões, resultando em redução de custos com medicamentos e melhoria da qualidade dos exames^{4,5}. Neste contexto, hipotetizamos que a plataforma, baseada em workflow sob medida para gestão de processos clínicos e administrativos, exerce um papel importante também na produtividade do serviço e na segurança dos pacientes. Neste trabalho apresentamos a experiência dos usuários da plataforma em relação à melhoria no gerenciamento de filas de atendimento e à segurança dos pacientes.

MÉTODOS

O software possui visões completas do protocolo de cada procedimento, expondo em um painel a etapa em que se encontra o paciente dentro da jornada do cuidado. A imagem abaixo mostra uma fila de pacientes com suas etapas e prioridades em tempo real. Cada usuário pode utilizar filtros específicos para visualizar apenas os pacientes pertencentes a determinada etapa do atendimento.



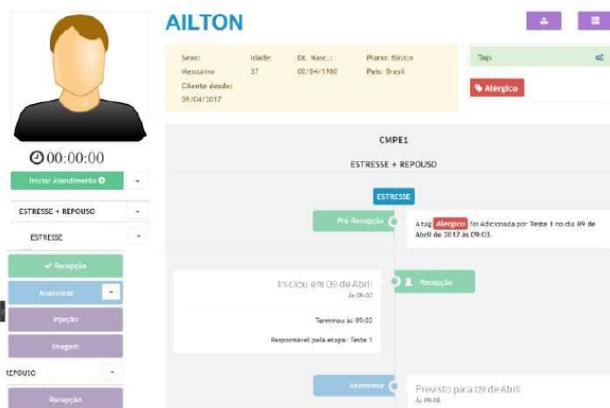
Etapa	Painel	Paciente	Masclado	Procedimento	Instituição	Operadora	Status
02/07/2019 16:00	Exame	TRABE ANAGAO VAIN	00000000	RM	CRM	BASTICELAR	Atendimento
03/07/2019 16:00	Exame	TESTE POSOLOGIA TC-PNH	00000000	RM	CRM	BASTICELAR	Atendimento
04/07/2019 16:00	Exame	INCIDENTE SUSTES	00000000	RM	CRM	BASTICELAR	Leads
06/07/2019 16:00	Exame	TESTE TABELT 1	00000000	FT	RET	BASTICELAR	Leads
06/07/2019 16:00	Exame	TESTE TABELT 1	00000000	FT	RET	JFFECO	Leads
07/07/2019 16:00	Exame	TESTE POSOLOGIA TFC	00000000	FT	RET	JFFECO	Atendimento
08/07/2019 16:00	Exame	TESTE TABELT 1	00000000	FT	RET	BASTICELAR	Atendimento
11/07/2019 16:00	Exame	TESTE TFC	00000000	FT	RET	JFFECO	Leads
11/07/2019 16:00	Exame	TESTE TFC	00000000	FT	RET	JFFECO	Atendimento

No exemplo a seguir, uma fila de infusão/ injeção de medicações com informações da posologia de acordo com o protocolo clínico e perfil do paciente.



St.	Nome do Paciente	Exame / Radiofarmaco	Massa Corpora	Regime de Atividade	Atividade prescrita	Hora Injeção	Atividade corrigida	Ações
1	APARECIDA DA SILVA PEPINO	Pet Onc / FLUORDEOIGLICOSE (F-18)	50 kg	500	5,00 mCi	05:30	13,93 mCi	Produzir
2	JOSE DOS SANTOS	Pet Onc / FLUORDEOIGLICOSE (F-18)	80 kg	500	8,00 mCi	06:20	26,97 mCi	Produzir
3	JULIA MARIA DOS SANTOS	Pet Onc / FLUORDEOIGLICOSE (F-18)	33 kg	500	3,30 mCi	06:30	13,46 mCi	Produzir

Por último, exemplificamos o workflow demonstrando todas as etapas e informações do paciente, possibilitando rastreabilidade de dados como medicações recebidas, etapas de atendimento, classificadores de risco, dentre outras.



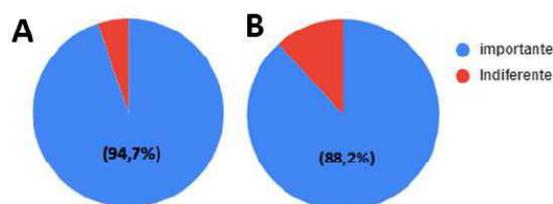
Este é um estudo qualitativo com aplicação de questionário para analisar o impacto do sistema Nuclearis na produtividade do serviço e na segurança dos pacientes. Para tal, utilizamos como métrica a percepção dos usuários em relação aos recursos de gerenciamento de filas e rastreabilidade do paciente.

O questionário utilizado é parte de um estudo contínuo de avaliação dos usuários do software, em que foram entrevistados 17 usuários até a presente data.

Os participantes da pesquisa foram profissionais médicos, tecnólogos, profissionais da qualidade e administrativos que trabalham com a solução nas seguintes instituições de saúde: Villela Pedras (Rio de Janeiro), Núcleos Radiologia e Medicina Nuclear (Brasília), Hospital Sírio Libanês (Brasília), Hospital Córdio Pulmonar (Salvador), Hospital São Rafael (Salvador), Grupo MND (Campinas).

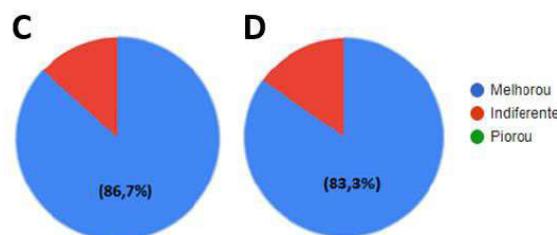
RESULTADOS

- A) Como você avalia a importância de um software para auxiliar na rastreabilidade do paciente?
 B) Como você avalia a importância de um software para gerenciar as filas de atendimento?



- C) Na sua experiência com outros sistemas, em relação ao Nuclearis, como você avalia o impacto na segurança do paciente para a administração dos radiofármacos?

- D) Na sua opinião, quanto o Nuclearis auxilia no gerenciando de filas no atendimento em relação a outros sistemas?



CONCLUSÕES

Na avaliação da maioria dos usuários, o software Nuclearis auxilia no gerenciamento de filas e na segurança do paciente, se mostrando uma ferramenta superior às outras disponíveis até o momento.

AGRADECIMENTOS

Ao CNPq, à FAPESB e à FINEP, pelo apoio ao desenvolvimento tecnológico da plataforma Nuclearis. Aos centros usuários do software e participantes da pesquisa.

REFERÊNCIAS

- [1] Ministério da Saúde. Ações e Programa: Programa Nacional de Segurança do Paciente.
- [2] Ministério da Saúde. Protocolos Básicos de Segurança do Paciente.
- [3] de Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ and Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review, 2008. *BMJ Quality & Safety* 17 (3 Suppl): 216-223.
- [4] Menezes V, Machado M, Queiroz C, Namias M, Reinande I, Vieira L, Gama J. How does Lean Six Sigma method improve healthcare practice in nuclear medicine departments? A successful case of dedicated software applications in oncological PET/CT. *J Nucl Med* 2018; 59 (1 Suppl): 1013.
- [5] Machado M, Menezes V, Queiroz C, Vigário A. et al. Inovação tecnológica para apoio à gestão e melhoria do cuidado. In: PRIMEIRO CONGRESSO DA REBRATS, 2019, Brasília. Anais eletrônicos. Campinas, Galoá, 2019.

ENDEREÇO PARA CORRESPONDÊNCIA

Email: dayane@nuclearis.com.br
 Celular: +55 (18) 997920078

Anexo B – Método automático para cálculo da composição corporal

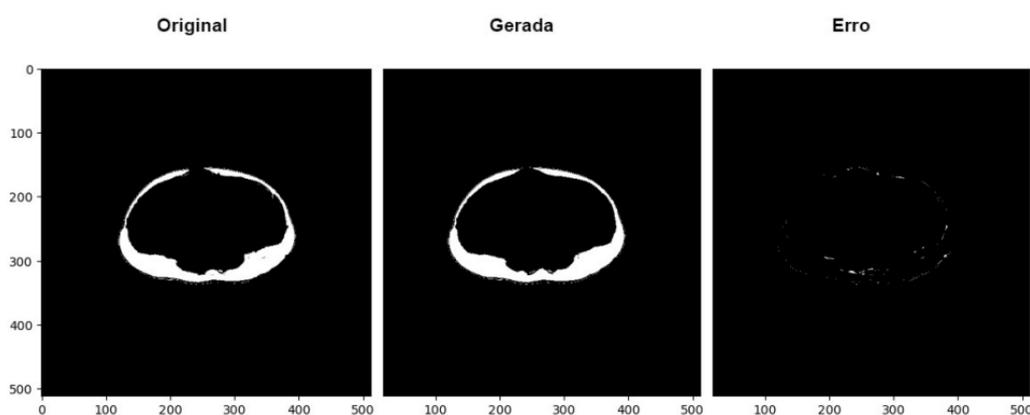
Marcos A. D. Machado^{1,2}, Bruno H. L. Anjos²

- 1) Departamento de Radiologia, Complexo Hospitalar Universitário Professor Edgard Santos/Universidade Federal da Bahia (UFBA)/ Ebserh, Salvador, Bahia, Brasil.
- 2) Departamento de Inteligência Artificial - Nuclearis Corporation, Recife, Pernambuco, Brasil

A composição corporal tem sido associada com desfechos de sobrevida em diferentes tipos de câncer (EBADI *et al*, 2020 & DA CUNHA JÚNIOR *et al*, 2021). A imagem de CT na terceira vértebra lombar foi proposta por DA CUNHA JÚNIOR e colegas (2021) para a análise da composição corporal de pacientes com mieloma múltiplo, servindo como biomarcador prognóstico.

Neste trabalho, utilizamos a coorte de pacientes com câncer de colo de útero para segmentação da gordura subcutânea de 47 pacientes, e outra coorte de 30 pacientes com outros tipos de câncer. A segmentação foi realizada utilizando a plataforma LIFEx para gerar os arquivos de aprendizado das segmentações em 10 a 20 cortes axiais dos 77 pacientes. Com isso, foi desenvolvida uma rede neural utilizando linguagem Python para segmentação automática da região de interesse para o cálculo da gordura subcutânea. Após a segmentação, o indicador de composição corporal é a média do número Hounsfield dos pixels. Utilizamos o Dice Score (F1) como métrica de desempenho do software de segmentação automática.

A figura abaixo demonstra o resultado do nosso software: à esquerda, a máscara de segmentação delineada pelo especialista, ao meio a máscara de segmentação resultante do nosso software, e à direita o erro entre as segmentações.



O resultado do Dice Score realizando o método *leave-one-out* foi de 0,95 (0,82-100), mostrando-se uma ferramenta viável para automação do cálculo de composição corporal.

Anexo C – Registro de *Software* no Instituto Nacional de Propriedade Industrial (INPI)

Os quatro *softwares* desenvolvidos neste trabalho foram depositados para registro no INPI, conforme segue:

- Sistema para gestão de serviços de saúde: gerenciamento de filas e automação do fluxo de pacientes

Número do processo: BR512023000174-0

- Método para particionamento de imagens gated-SPECT

Número do processo: BR512023000147-3

- Cálculo automático da composição corporal através da tomografia computadorizada

Número do processo: BR512023000125-2

- Sistema de visualização e análise de imagens médicas com inteligência artificial para diagnóstico da COVID-19

Número do processo: BR512023000157-0



20/01/2023 870230005811
20:05

29409191960882040

Pedido de Registro de Programa de Computador - RPC

Número do Processo: 512023000174-0

Dados do Titular

Titular 1 de 1

Nome ou Razão Social: NUCLEARIS SISTEMAS EM MEDICINA NUCLEAR LTDA

Dados do Programa

Data de Publicação: 10/01/2021

Data de Criação: 30/12/2020

- § 2º do art. 2º da Lei 9.609/98: "Fica assegurada a tutela dos direitos relativos a programa de computador pelo prazo de cinquenta anos contados a partir de 1º de janeiro do ano subsequente ao da sua publicação ou, na ausência desta, da sua criação"

Título: Sistema para gestão de serviços de saúde: gerenciamento de filas e automação do fluxo de pacientes

Algoritmo hash: SHA-512 - Secure Hash Algorithm

Resumo digital hash: 32b97e196d4677fb0a73526bda21eae7daff35c6bbb8cffd256bd60738
64b47d99048b5992b1ceb10c3ad7899d7612673148fb27dd45b71fd
6639c5c24edcb1

18/01/2023 870230004920
13:49

29409191960704606

Pedido de Registro de Programa de Computador - RPC

Número do Processo: 512023000147-3

Dados do Titular

Nome ou Razão Social: RADTEC SERVICOS EM FISICA MEDICA LTDA

Dados do Programa

Data de Publicação: 17/12/2021

Data de Criação: 15/03/2021

- § 2º do art. 2º da Lei 9.609/98: "Fica assegurada a tutela dos direitos relativos a programa de computador pelo prazo de cinquenta anos contados a partir de 1º de janeiro do ano subsequente ao da sua publicação ou, na ausência desta, da sua criação"

Título: Método para particionamento de imagens gated-SPECT

Algoritmo hash: SHA-512 - Secure Hash Algorithm

Resumo digital hash: 58a622c61f68f334bbea9ccbe547f39cdc1af2c8363a78f278960ac10a
5ec06c00312d2ac77b677f6f1139ab686a25d260322f43ec4dd15b942
19478d109e75a



16/01/2023

870230004334

17:22



29409191960487265

Pedido de Registro de Programa de Computador - RPC

Número do Processo: 512023000125-2

Dados do Titular

Titular 1 de 1

Nome ou Razão Social: NUCLEARIS SISTEMAS EM MEDICINA NUCLEAR LTDA

Dados do Programa

Data de Publicação: 11/01/2023

Data de Criação: 20/12/2022

- § 2º do art. 2º da Lei 9.609/98: "Fica assegurada a tutela dos direitos relativos a programa de computador pelo prazo de cinquenta anos contados a partir de 1º de janeiro do ano subseqüente ao da sua publicação ou, na ausência desta, da sua criação"

Título: Cálculo automático da composição corporal através da tomografia computadorizada

Algoritmo hash: SHA-512 - Secure Hash Algorithm

Resumo digital hash: f48ecb5c54e4f4b59eaa1920b3dbe326dbad5392bea8c3562b5587cc
2b0a29ffa501fac563cfebd84762400ba816d4bf6a1a11fe6eb57a0253
927d9a2ef6826e

19/01/2023 870230005554
17:33

29409191960783298

Pedido de Registro de Programa de Computador - RPC

Número do Processo: 512023000157-0

Dados do Titular

Titular 1 de 1

Nome ou Razão Social: NUCLEARIS SISTEMAS EM MEDICINA NUCLEAR LTDA

Dados do Programa

Data de Publicação: 30/01/2022

Data de Criação: 30/11/2021

- § 2º do art. 2º da Lei 9.609/98: "Fica assegurada a tutela dos direitos relativos a programa de computador pelo prazo de cinquenta anos contados a partir de 1º de janeiro do ano subsequente ao da sua publicação ou, na ausência desta, da sua criação"

Título: Sistema de visualização e análise de imagens médicas com inteligência artificial para diagnóstico da COVID-19

Algoritmo hash: SHA-512 - Secure Hash Algorithm

Resumo digital hash: 1889179cdbad07243ae0772a21253bf8eec40d77749a0301d90015c5
a0662e692b36e783c27d456ab6a93b1aa7ee91044032c07f9edb96df
575c0a8520f5779a

Anexo D - Pareceres do Comitê de Ética

HOSPITAL SÃO
RAFAEL/MONTE TABOR-BA



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Análise da textura de tumores em imagens PET/CT como biomarcador diagnóstico e prognóstico.

Pesquisador: Marcos Antônio Dórea Machado

Área Temática:

Versão: 1

CAAE: 79258017.6.0000.0048

Instituição Proponente: Hospital São Rafael/Monte Tabor-BA

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 2.381.285

UFPE - HOSPITAL DAS
CLÍNICAS DA UNIVERSIDADE
FEDERAL DE PERNAMBUCO -
HC/UFPE



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: RADIÔMICA COMO BIOMARCADOR PROGNÓSTICO DO 18F-FDG PET/CT NO CÂNCER DE COLO UTERINO

Pesquisador: NADJA ROLIM GONCALVES DE ALENCAR

Área Temática:

Versão: 2

CAAE: 23650619.5.0000.8807

Instituição Proponente: EMPRESA BRASILEIRA DE SERVICOS HOSPITALARES - EBSERH

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 3.780.319

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



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

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



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Biobancos digitais: plataforma para gestão epidemiológica e medicina personalizada

Pesquisador: Marcos Antônio Dórea Machado

Área Temática:

Versão: 1

CAAE: 58972522.7.0000.0049

Instituição Proponente: UNIVERSIDADE FEDERAL DA BAHIA

Patrocinador Principal: Financiadora de Estudos e Projetos - FINEP

DADOS DO PARECER

Número do Parecer: 5.462.552

Anexo E – Artigos submetidos (comprovantes)

Caro (a) Marcos Machado,

Agradecemos a submissão do manuscrito "Prediction of Treatment Failure in Hodgkin Lymphoma: A Machine Learning Radiomic Approach in Baseline 18F-FDG PET/CT" para a revista Revista Brasileira de Física Médica.

Acompanhe o progresso da sua submissão por meio da interface de administração do sistema, disponível em:

URL da submissão: <https://www.rbfm.org.br/rbfm/authorDashboard/submission/680>

Login: doreamachado

Em caso de dúvidas, entre em contato via e-mail.

Agradecemos por considerar nossa revista como meio de compartilhar seu trabalho.

Cordialmente,

Ana Maria Marques da Silva

Dear Mr Machado,

We have received the reports from our advisors on your manuscript, "Multi-center Integrating Radiomics, Structured Reports and Machine Learning Algorithms for Assisted Classification of COVID-19 in Lung Computed Tomography", which you submitted to Journal of Medical and Biological Engineering.

Based on the advice received, I have decided that your manuscript could be reconsidered for publication should you be prepared to incorporate major revisions. When preparing your revised manuscript, you are asked to carefully consider the reviewer comments which can be found below, and submit a list of responses to the comments. You are kindly requested to also check the website for possible reviewer attachment(s).

In order to submit your revised manuscript, please access the journal's Editorial Manager site.

Multi-center Integrating Radiomics, Structured Reports and Machine Learning Algorithms for Assisted Classification of COVID-19 in Lung Computed Tomography

We thank the editor and reviewer for the comments and for considering our work for publication.

Reviewer #1:

1. SARS-CoV-2 causes the ground glass opacity of the lung. How can this image be distinguished from other cancerous or inflammatory lesions?

Thank you for this important comment. We improved the discussion to clarify this issue:

"The CT aspects of COVID-19 lesions are usually presented with ground glass opacity, rounded or polygonal, and its distribution is mostly peripheral and bilateral, without pulmonary consolidation (table 2). Because inflammatory lesions in the lungs are not pathognomonic, the etymology of some pneumonia cases are tough to differentiate using only CT without complementary technologies".

2. Does too few subjects affect the ratio of sensitivity and specificity?

Sensitivity and specificity ratios are not affected by the sample size, but the confidence interval in larger samples is smaller. Our results are presented with 95% confidence interval and 98% statistical power.

Reference: OpenEpi software [18].

3. Have you ever used AI to differentiate CT images of other cancerous or inflammatory lesions in the way of blindly inferring?

Our AI model was trained/ validated using subjects affected by pneumonia. Controls were selected using criteria of pneumonia diagnosis, imaged during January-2018 and October-2019. On the other hand, cases were selected between April-2020 and April-2021, also based on the criteria of pneumonia diagnosis, but with confirmed RT-PCR examination. Therefore, our AI models do not have capability to differentiate cancerous or inflammatory lesions, but it only informs how likely the lesions' etymology is COVID-19 or not. |