

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS APLICADAS À SAÚDE DO
ADULTO

JÚLIO GUERRA DOMINGUES

Análise e desenvolvimento de redes neurais convolucionais para o diagnóstico radiográfico de osteoartrite dos joelhos no Estudo Longitudinal da Saúde do Adulto – Musculoesquelético (ELSA-Brasil MSK)

BELO HORIZONTE

2022

JÚLIO GUERRA DOMINGUES

Análise e desenvolvimento de redes neurais convolucionais para o diagnóstico radiográfico de osteoartrite dos joelhos no Estudo Longitudinal da Saúde do Adulto – Musculoesquelético (ELSA-Brasil MSK)

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Aplicadas à Saúde do Adulto da Faculdade de Medicina da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Mestre.

Orientadora: Prof^a. Dr^a. Rosa Weiss Telles

Co-orientadores:

Prof^a. Dr^a. Sandhi Maria Barreto

Prof. Dr. Alexei Manso Corrêa Machado

BELO HORIZONTE

2022

D671a Domingues, Júlio Guerra.
Análise e desenvolvimento de Redes Neurais Convolucionais para o diagnóstico radiográfico de Osteoartrite dos joelhos no Estudo Longitudinal da Saúde do Adulto – Musculoesquelético (ELSA-Brasil MSK) [recursos eletrônicos]. / Júlio Guerra Domingues. - - Belo Horizonte: 2022.
97f.: il.
Formato: PDF.
Requisitos do Sistema: Adobe Digital Editions.

Orientador (a): Rosa Weiss Telles.
Coorientador (a): Sandhi Maria Barreto; Alexei Manso Corrêa Machado.
Área de concentração: Doenças Crônico-Degenerativas.
Dissertação (mestrado): Universidade Federal de Minas Gerais, Faculdade de Medicina.

1. Osteoartrite do Joelho. 2. Redes Neurais de Computação. 3. Aprendizado de Máquina. 4. Diagnóstico por Computador. 5. Estudos Epidemiológicos. 6. Dissertação Acadêmica. I. Telles, Rosa Weiss. II. Barreto, Sandhi Maria. III. Machado, Alexei Manso Corrêa. IV. Universidade Federal de Minas Gerais, Faculdade de Medicina. V. Título.

NLM: WE 870

Bibliotecário responsável: Fabian Rodrigo dos Santos CRB-6/2697



UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS APLICADAS À SAÚDE DO ADULTO

FOLHA DE APROVAÇÃO

ANÁLISE E DESENVOLVIMENTO DE REDES NEURAI CONVOLUCIONAIS PARA O DIAGNÓSTICO RADIOGRÁFICO DE OSTEOARTRITE DOS JOELHOS NO ESTUDO LONGITUDINAL DA SAÚDE DO ADULTO – MUSCULOESQUELÉTICO (ELSA-BRASIL MSK)

JÚLIO GUERRA DOMINGUES

Dissertação de Mestrado defendida e aprovada, no dia vinte e dois de novembro de dois mil e vinte e dois, pela Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Ciências Aplicadas à Saúde do Adulto da Universidade Federal de Minas Gerais constituída pelos seguintes professores doutores:

Rosa Weiss Telles - Orientadora
UFMG

Sandhi Maria Barreto - Coorientadora
UFMG

Alexei Manso Correa Machado - Coorientador
UFMG

Luciana Costa Silva
UFMG

Flavio Vinicius Diniz de Figueiredo
UFMG

Belo Horizonte, 22 de novembro de 2022.



Documento assinado eletronicamente por **Rosa Weiss Telles, Professora do Magistério Superior**, em 01/12/2022, às 14:24, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Alexei Manso Correa Machado, Servidor(a)**, em 01/12/2022, às 15:11, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Flavio Vinicius Diniz de Figueiredo, Professor do Magistério Superior**, em 05/12/2022, às 10:04, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Sandhi Maria Barreto, Membro de comissão**, em 13/12/2022, às 12:24, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Luciana Costa Silva, Usuário Externo**, em 29/12/2022, às 17:15, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1856239** e o código CRC **D72AAFE6**.

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Reitora: Professora Sandra Regina Goulart Almeida

Vice-Reitor: Professor Alessandro Fernandes Moreira

Pró-Reitora De Pós-Graduação: Professora Isabela Almeida Pordeus

Pró-Reitor De Pesquisa: Professor Fernando Marcos Dos Reis

FACULDADE DE MEDICINA

Diretora Da Faculdade De Medicina: Professora Alamanda Kfoury Pereira

Vice-Diretora Da Faculdade De Medicina: Professora Cristina Gonçalves Alvim

Coordenador Do Centro De Pós-Graduação: Professor Tarcizo Afonso Nunes

Subcoordenadora Do Centro De Pós-Graduação: Professora Eli lola Gurgel Andrade

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS APLICADAS À SAÚDE DO ADULTO

Coordenadora do Programa de Pós-Graduação em Ciências Aplicadas à Saúde do Adulto: Professora Teresa Cristina De Abreu Ferrari

Subcoordenadora do Programa de Pós-Graduação em Ciências Aplicadas à Saúde do Adulto: Professora Luciana Costa Faria

Colegiado do Programa de Pós-Graduação em Ciências Aplicadas à Saúde do Adulto:
Professora Claudia Alves Couto

Professora Gilda Aparecida Ferreira

Professora Karina Braga Gomes Borges

Professora Luciana Costa Faria

Professora Luciana Diniz Silva

Professora Melissa Orlandin Premaor

Professora Teresa Cristina De Abreu Ferrari

À Luciana, meu amor e maior exemplo.

Aos nossos filhos, Alice e João, pela inspiração.

Aos meus pais e irmãos, pelo apoio desde sempre.

*"The mind that opens to new ideas never
returns to its original size."*

(Albert Einstein)

AGRADECIMENTOS

Agradeço à minha esposa, Luciana, pela paciência ao escutar tantas ideias em efervescência e pela sabedoria ao me aconselhar. Agradeço também pelas diversas sugestões e revisões, sem as quais não teria alcançado o presente trabalho.

À minha orientadora Rosa Weiss Telles, pelo exemplo de profissionalismo e seriedade. Por me guiar em momentos de dúvida e por me deixar fazer parte da equipe do ELSA-Brasil MSK.

Aos meus co-orientadores, Sandhi e Alexei. À Sandhi pela clareza e resolutividade ao lidar com assuntos tão complexos e ao Alexei pela perspectiva computacional e comentários sempre precisos.

Agradeço ao Adriano Veloso pela profícua parceria com o Departamento de Ciência da Computação e por possibilitar os recursos para o desenvolvimento do trabalho. À Daniella, pela grande contribuição na concepção da rede neural e disponibilidade.

Aos radiologistas do Hospital Madre Teresa pelos ensinamentos e fundamental contribuição para a minha formação.

Agradeço à Luciana Costa por ter iniciado o caminho que trilhei no ELSA-Brasil MSK e à Jesiana por ter me apresentado ao ELSA-Brasil.

A todos os profissionais do ELSA-Brasil e ELSA-Brasil MSK, pela excelência no trabalho realizado, que permitiu o desenvolvimento do presente volume.

Agradeço, por fim, a todos os participantes do ELSA-Brasil, especialmente do Centro de Investigação Minas Gerais, pelo altruísmo, que possibilita o nosso aprendizado contínuo.

RESUMO

Introdução: A osteoartrite dos joelhos (OAJ) é usualmente diagnosticada por meio da interpretação de radiografias utilizando-se o sistema de classificação de Kellgren-Lawrence (KL), que envolve a avaliação de osteófitos e redução de espaço articular. Em estudos epidemiológicos longitudinais de larga escala, tal classificação é realizada por médicos especialistas, em um processo laborioso e demorado que exige treinamento rigoroso. Nesse contexto, a aplicação de modelos computacionais para classificação automática da OAJ pode contribuir para o fluxo de diagnóstico, reduzindo o número total de exames a serem avaliados por humanos. Dentre os modelos de classificação automática, destacam-se as redes neurais convolucionais (RNC), que têm demonstrado resultados promissores na área de diagnóstico médico por imagem. Estudos recentes demonstraram, todavia, que o desenvolvimento e a validação das RNC atualmente existentes estão concentrados na população de países desenvolvidos, demonstrando a necessidade de maior validação externa de tais ferramentas. **Objetivos:** Avaliar o desempenho de uma RNC já publicada para o diagnóstico de OAJ nas radiografias dos joelhos dos participantes do ELSA-Brasil MSK e desenvolver uma nova RNC treinada a partir dos exames da linha de base do ELSA-Brasil MSK. **Método:** Trata-se de um estudo transversal realizado com exames radiográficos da linha de base do ELSA-Brasil MSK. O ELSA-Brasil MSK, estudo ancilar ao Estudo Longitudinal da Saúde do Adulto (ELSA-Brasil), incluiu, em sua linha de base, 2901 servidores públicos ativos ou aposentados de duas grandes instituições de ensino e pesquisa de Minas Gerais. Para a verificação de desempenho da RNC previamente publicada, foi selecionada uma amostra de conveniência composta por 244 radiografias. As predições da RNC foram comparadas com a classificação KL e diagnóstico de OA realizada pelos radiologistas do ELSA-Brasil MSK. Para o desenvolvimento da RNC, foram utilizadas todas as radiografias da linha de base do ELSA-Brasil MSK (5660 joelhos). Os exames foram interpretados por médico radiologista com treinamento específico e calibração já publicada. **Resultados:** A RNC previamente publicada apresentou área sob a curva característica de operação do receptor (AUC) de 0,901 para o diagnóstico de OA (IC 95% 0,858-0,945). A RNC desenvolvida a partir das radiografias e classificações do ELSA-Brasil MSK apresentou (AUC) de 0,866 (IC 95% 0,842-0,882). **Conclusões:** As RNC avaliadas apresentaram desempenho equivalente para a avaliação das radiografias do ELSA-

Brasil MSK. A utilização de tais programas, isolada ou conjuntamente, pode auxiliar como mecanismo de triagem, reduzindo o número total de exames avaliados pelos radiologistas do estudo, e/ou como ferramenta de segunda leitura, contribuindo com a redução de possíveis erros de interpretação.

Palavras-chave: Osteoartrite do Joelho; Radiografia; Redes Neurais de Computação; Aprendizado de Máquina; Diagnóstico por Computador; Estudos Epidemiológicos.

ABSTRACT

Introduction: Knee Osteoarthritis (KOA) is frequently diagnosed by evaluating radiographies according to the Kellgren-Lawrence (KL) grading system, which involves the assessment of osteophytes and articular joint-space narrowing. In large-scale longitudinal epidemiological studies, such classification is performed by specialist physicians, in a laborious and time-consuming process that requires rigorous training. In this context, the application of computational models for automated KOA classification can contribute to the diagnostic flow, reducing the total number of exams to be evaluated by humans. Among the automatic classification models, convolutional neural networks (CNN) stand out, having demonstrated promising results in the medical imaging diagnosis field. Recent studies have shown, however, that the development and validation of currently existing CNN is concentrated in the population of developed countries, demonstrating the need for greater external validation of such methods. **Objectives:** To evaluate the performance of a previously published CNN for KOA diagnosis on radiographs of ELSA-Brasil MSK participants and to develop a CNN from scratch trained with the ELSA-Brasil MSK baseline exams. **Method:** This is a cross-sectional study carried out with radiographies from ELSA-Brasil MSK's baseline. The ELSA-Brasil MSK, an ancillary study to the Longitudinal Study of Adult Health (ELSA-Brasil), included, in its baseline, 2901 active or retired public servants from two large teaching and research institutions in Minas Gerais. To verify the performance of the previously published CNN, a convenience sample consisting of 243 radiographs was selected. CNN predictions were compared with the KL classification and OA diagnosis performed by ELSA-Brasil MSK's radiologists. For the development of the CNN, all radiographies from the baseline of the ELSA-Brasil MSK (5660 knees) were used. The exams were interpreted by a radiologist with specific training and calibration, previously published. **Results:** The previously published CNN presented an area under receptor operating characteristics curve (AUC) of 0.901 for the diagnosis of OA (CI 95%: 0.858-0.945). The CNN developed from the radiographs and classifications of the ELSA-Brasil MSK presented an (AUC) of 0.866 (95% CI: 0.842-0.882). **Conclusions:** The evaluated CNN presented similar performance for the evaluation of the ELSA-Brasil MSK radiographies. The use of such models can help as a screening tool, reducing the total number of exams evaluated by the radiologists of the

study, and/or as a dual reading tool, contributing to the reduction of possible interpretation errors.

Keywords: Osteoarthritis, Knee; Radiography; Neural Networks, Computer; Machine Learning; Diagnosis, Computer-Assisted; Epidemiologic Studies.

I. Lista de abreviaturas e siglas

AB – Acurácia balanceada

ABP – Acurácia balanceada ponderada (do inglês *weighted balanced accuracy*)

ACD – Diagnóstico computacional automatizado (*automated computer diagnosis*)

ACR – Colégio Americano de Reumatologia (*American College of Rheumatology*)

AM – Aprendizado de máquina (*machine learning*)

AP – Aprendizado profundo (*deep learning*)

AUC – Área sob a curva característica de operação do receptor (*Area Under the Receiver Operating Characteristics Curve*)

CAD – Diagnóstico auxiliado por computador (*computer aided diagnosis*)

CEFET-MG – Centro Federal de Educação Tecnológica de Minas Gerais

CI-MG – Centro de Investigação de Minas Gerais

CEP – Comitê de Ética em Pesquisa

CHECK – *Cohort Hip and Cohort Knee study*

CNN – *Convolutional Neural Network*

CONEP – Comissão Nacional de Ética em Pesquisa

CT – Computed tomography

DL – *Deep learning*,

ELSA-Brasil – Estudo Longitudinal da Saúde do Adulto

ELSA-Brasil MSK – Estudo Longitudinal da Saúde do Adulto - Musculoesquelético

EUA – Estados Unidos da América

EULAR – Liga Europeia Contra o Reumatismo (*European League Against Rheumatism*)

FHS – *Framingham Heart Study*

FOS – *Framingham Osteoarthritis Study*

IA – Inteligência artificial (*artificial intelligence*)

ICD-9-CM – 9ª edição da Classificação Internacional de Doenças

KL – Kellgren-Lawrence

MEPS-HC – Componente domiciliar da pesquisa do painel de despesas médicas dos EUA (*Household Component - Medical Expenditure Panel Survey*)

ML – *Machine learning*

MOST – *Multi-Centre Osteoarthritis Study*

MRI – *Magnetic resonance imaging*

NICE – Instituto Nacional de Saúde e Excelência em Cuidados do Reino Unido (*National Institute for Health and Care Excellence*)

OA – Osteoartrite (*osteoarthritis*)

OAI – *Osteoarthritis Initiative*

OARSI – *Osteoarthritis Research Society International*

PA – Posteroanterior

PET – Tomografia por emissão de pósitrons (*positron emission tomography*)

PPG CASA-UFMG – Programa de Pós-Graduação em Ciências Aplicadas à Saúde do Adulto da Universidade Federal de Minas Gerais (PPG CASA-UFMG).

RM – Ressonância magnética

RNA – Rede neural artificial

RNC – Rede neural convolucional (*convolutional neural network*)

RX – Radiografia

SPECT – Tomografia computadorizada por emissão de fóton único (*single photon emission computed tomography*)

TC – Tomografia computadorizada

TCLE – Termo de Consentimento Livre e Esclarecido

UFMG – Universidade Federal de Minas Gerais

US – Ultrassonografia

II. Lista de tabelas

Tabela 1: Performance dos principais achados do exame físico e radiográficos para o diagnóstico de osteoartrite dos joelhos 27

Artigo 1

Table 1: Confusion matrix between CNN's predicted KL classifications related to ground-truth..... 54

Table 2: Confusion matrix between CNN's OA predictions and ground-truth 54

Artigo 2

Tabela 1: Valores de acurácia, acurácia balanceada, acurácia balanceada ponderada, sensibilidade, especificidade, precisão e escores F1 e F2 para cada limiar definido. 70

III. Lista de figuras

Figura 1: Custos diretos, indiretos e intangíveis relacionados à osteoartrite	24
Figura 2: A história natural da osteoartrite e seus biomarcadores.....	25
Figura 3: Exemplos de radiografias dos joelhos com diferentes classificações KL. .	32
Figura 4: Sumário dos diferentes algoritmos utilizados em aprendizado de máquina	40
Figura 5: Diagrama de Venn empilhado demonstra a relação entre as áreas da Inteligência Artificial.....	42
Figura 6: Fluxograma dos participantes incluídos na linha de base do Estudo Longitudinal de Saúde do Adulto – Musculoesquelético (ELSA-Brasil MSK).	46

Artigo 1

Figure 1: False negatives examples (three different subjects: A, B and C). Femoral definite osteophytes (red arrows) were not considered in CNN's reading..	55
Figure 2: False positive example: both radiologists considered this a normal exam (KL 0), while CNN classified it as KL 2..	56

Artigo 2

Figura 1: Ilustração esquemática da arquitetura DenseNet. Os pares de camadas são conectados, permitindo que elementos das primeiras camadas (como bordas) possam ser também utilizados nas últimas	66
Figura 2: Demonstração do pré-processamento.....	67
Figura 3: Curva característica de operação do receptor do modelo.	69

SUMÁRIO

1	CONSIDERAÇÕES INICIAIS	18
1.1	Apresentação	18
1.2	Antecedentes científicos	20
1.2.1	Osteoartrite.....	20
1.2.2	Osteoartrite de joelhos	21
1.2.3	Estudos longitudinais sobre Osteoartrite	37
1.2.4	Inteligência artificial	39
2	OBJETIVOS	45
2.1	Objetivos específicos	45
3	MÉTODOS	46
3.1	Desenho e Amostra	46
3.2	Aspectos éticos	47
3.3	Exame Radiográfico	47
3.4	Interpretação humana	47
4	RESULTADOS	49
4.1	Artigo original 1	49
4.2	Artigo original 2	61
5	CONSIDERAÇÕES FINAIS	80
6	REFERÊNCIAS BIBLIOGRÁFICAS	82
7	APÊNDICES	90
I.	Acesso ao código do modelo desenvolvido (Kneelsa)	90
8	ANEXOS	91
I.	Trabalho apresentado no 51º Congresso Brasileiro de Radiologia e Diagnóstico por Imagem	91
II.	Trabalho apresentado no British Institute of Radiology Annual Congress 2022	95
III.	Folha de Aprovação do Comitê de Ética	96

1 CONSIDERAÇÕES INICIAIS

1.1 Apresentação

O presente volume trata de dissertação de mestrado desenvolvida no Programa de Pós-Graduação em Ciências Aplicadas à Saúde do Adulto da Universidade Federal de Minas Gerais (PPG CASA-UFMG). Este trabalho está inserido na linha de pesquisa “Doenças crônico-degenerativas”, integrando o “Estudo Longitudinal de Saúde do Adulto – Musculoesquelético (ELSA-Brasil MSK)”.

O Estudo Longitudinal de Saúde do Adulto (ELSA-Brasil) é uma coorte multicêntrica iniciada em 2008 que acompanha servidores públicos, ativos e aposentados, em seis estados brasileiros: Bahia, Espírito Santo, Minas Gerais, Rio de Janeiro, Rio Grande do Sul e São Paulo (SCHMIDT et al., 2015). O ELSA-Brasil visa à investigação da incidência e progressão de doenças crônicas não-transmissíveis, examinando os fatores biológicos, comportamentais, ambientais, ocupacionais, psicológicos e sociais a elas relacionados (AQUINO et al., 2012).

Um dos estudos ancilares do ELSA-Brasil é o ELSA-Brasil MSK, desenvolvido no Centro de Investigação de Minas Gerais (CI-MG) e iniciado em 2012. Este é composto, em sua linha de base, por 2901 servidores ativos e aposentados, com idade entre 38 e 79 anos (média 56,0, desvio-padrão 8,9, 52,9% mulheres), da Universidade Federal de Minas Gerais (UFMG) e do Centro Federal de Educação Tecnológica de Minas Gerais, Campus Belo Horizonte (CEFET-MG) (TELLES et al., 2022). As avaliações do ELSA-Brasil MSK aconteceram em 2 ondas (2012-2014 e 2016-2018), estando em andamento a 3ª onda iniciada em agosto de 2022.

O ELSA-Brasil é financiado pelo Ministério da Saúde (Departamento de Ciência e Tecnologia) e pelo Ministério de Ciência e Tecnologia (Financiadora de Estudos e Projetos e CNPq-BR) do Brasil, 01 10 0742-00 BA, 01 12 0284-00 ES, 01 10 0746-00 MG, 01 11 0093-01 RJ, 01 10 0643-03 RS, 01 10 0773-00 SP. O ELSA-Brasil MSK recebe apoio e agradece o recebimento de fomento dos seguintes órgãos/fundações: CAPES-BR (post-doc/SUS 054/2010), FAPEMIG-BR (APQ-00921-16, APQ-00549-22) e CNPq-BR (423585/2016-9, 404728/2021-9).

O ELSA-Brasil foi aprovado pela Comissão Nacional de Ética em Pesquisa (CONEP) [protocolo 976/2006] e pelos Comitês de Ética em Pesquisa (CEP) de todas as instituições participantes. O estudo ancilar ELSA-Brasil MSK, foi aprovado por meio de emendas submetidas ao CEP da UFMG (CAAE 0186.1.203.000-06 e 47125015.4.1001.5149).

Este volume de dissertação está apresentado no formato em artigo, de acordo com o regulamento do PPG CASA-UFMG, e preenche requisito parcial para a obtenção do título de mestre em Ciências Aplicadas à Saúde do Adulto. O volume contém:

- I. Considerações iniciais
 - a. Apresentação
 - b. Antecedentes científicos
- II. Objetivos geral e específicos
- III. Métodos
- IV. Resultados
 - a. Artigo original 1
 - b. Artigo original 2
- V. Considerações finais
- VI. Referências bibliográficas
- VII. Apêndices
 - a. Acesso ao código do modelo desenvolvido (Kneelsa)
- VIII. Anexos
 - a. Trabalho apresentado no 51º Congresso Brasileiro de Radiologia e Diagnóstico por Imagem
 - b. Trabalho apresentado no *British Institute of Radiology Annual Congress 2022*
 - c. Folha de Aprovação do Comitê de Ética em Pesquisa da UFMG

1.2 Antecedentes científicos

1.2.1 Osteoartrite

Distúrbios do sistema musculoesquelético afetam diretamente a saúde e a qualidade de vida dos indivíduos, estando entre as principais causas de anos vividos com incapacidade em todo o mundo (ABBAFATI et al., 2020; JIN et al., 2021). Dentre esses distúrbios, destaca-se a osteoartrite (OA), forma mais comum de artrite e um dos diagnósticos mais frequentes da prática médica (SANTO; OKEYODE, 2018). Estima-se que ela afete 303 milhões de pessoas em todo mundo (SAFIRI et al., 2020), estando entre as cinco maiores causas de anos de vida perdidos ajustados por incapacidade (do inglês *Disability-Adjusted Life-Years* – DALYs) (MURRAY et al., 2015).

A estimativa da prevalência de OA apresenta acentuada variabilidade entre diferentes estudos. Isso se deve a diferenças nas populações avaliadas (ex.: idade, nacionalidade) e articulações acometidas, fonte dos dados e definições utilizadas para o diagnóstico de OA (ALLEN; THOMA; GOLIGHTLY, 2022).

Cisternas et al. (2016) propuseram a utilização dos seguintes conceitos de provável OA para a estimativa da prevalência da doença, com base no componente domiciliar da pesquisa do painel de despesas médicas dos EUA (MEPS-HC): código 715 da Modificação Clínica da 9ª edição da Classificação Internacional de Doenças (ICD-9-CM) OU códigos 716 ou 719 da ICD-9-CM associados ao diagnóstico prévio de artrite, exceto reumatoide, autorreferido (questionário). De acordo com tais critérios, estimaram a prevalência da OA (em qualquer articulação) em 30,8 milhões de pessoas nos Estados Unidos (13,4% da população). Tal estimativa está em consonância com outros estudos epidemiológicos mais recentes, que avaliam a prevalência entre 10,7% e 29,3% em diferentes países (ALLEN; THOMA; GOLIGHTLY, 2022). Estudos demonstram elevações em prevalência, incidência e anos vividos com incapacidade ajustados à idade de 8% a 10% desde 1990 (SAFIRI et al., 2020).

A OA afeta principalmente as articulações que sustentam o peso do corpo, como os joelhos e quadris, e tem sido entendida como uma doença que envolve todas as

estruturas que compõem a articulação, e não mais apenas ossos e cartilagem (BIJLSMA; BERENBAUM; LAFEBER, 2011; NEVITT; FELSON; LESTER, 2017).

1.2.2 Osteoartrite de joelhos

O acometimento dos joelhos destaca-se pela alta frequência e pela disfunção e incapacidade relacionadas, sendo a afecção responsável pela maior parcela da incapacidade observada em pacientes idosos, estando à frente de problemas de saúde relacionados a altos índices de morbidade nessa população, como as doenças cardiovasculares, o diabetes e a fratura de quadril (VOS et al., 2012).

1.2.2.1 Epidemiologia

A incidência de OA dos joelhos (OAJ) nos EUA, em diferentes estudos, foi estimada entre 80 e 240/100.000 pessoas-ano, sendo mais incidente em mulheres, com aumento significativo a partir dos 50 anos de idade (NEOGI; ZHANG, 2013; OLIVERIA et al., 1995). OA radiográfica foi encontrada em 19% dos participantes do estudo Framingham com idade maior ou igual a 45 anos (FELSON et al., 1995) e, estima-se, teve sua prevalência duplicada desde meados do século XX (WALLACE et al., 2017).

Estudos recentes evidenciam prevalência de OAJ sintomática de 13,8% na Espanha, 17,4% no Reino Unido, 14,6% na China (BLANCO et al., 2021; LI et al., 2020; PEAT et al., 2020). O risco de desenvolver OAJ sintomática na população geral foi estimado entre 14% e 45% (LOSINA et al., 2013; MURPHY et al., 2008).

Poucos estudos estimaram a prevalência de OAJ na população brasileira. Telles e et al. (2022) identificaram OAJ radiográfica em 18,1% no maior estudo longitudinal para estudo de doenças musculoesqueléticas do Brasil.

1.2.2.2 Fatores de risco

Diversos fatores foram associados a risco aumentado de desenvolvimento de OAJ. Uma revisão sistemática de Silverwood et al. (2015) estimou maiores frações atribuíveis populacionais a sobrepeso, obesidade e lesões articulares prévias. O sexo feminino foi também considerado um fator de risco importante. A idade é considerada um importante fator de risco, não só devido à exposição a outros diferentes fatores de

risco ao longo da vida, mas também pelo desenvolvimento de alterações degenerativas articulares (ZHANG; JORDAN, 2010).

Fatores ergonômicos/ocupacionais (por ex.: carregamento de peso e agachamentos frequentes) e desalinhamentos articulares (por ex.: desvio em varo ou valgo) apresentam-se como fatores de risco de menor magnitude (SILVERWOOD et al., 2015). Não há evidências de que atividades físicas recreacionais reduzam ou aumentem o risco de OAJ (BARBOUR et al., 2014; FELSON et al., 2007), apesar de alguns esportes de alto impacto contribuírem para maior exposição a lesões articulares traumáticas que, por sua vez, estão relacionadas ao aumento do risco de desenvolvimento de OA (DRIBAN et al., 2017; FREIBERG; BOLM-AUDORFF; SEIDLER, 2021).

A OAJ compartilha fatores de risco com diversas doenças crônicas não transmissíveis. Yoshimura et al. (2012) demonstraram que o acúmulo de componentes da síndrome metabólica (como sobrepeso, hipertensão, dislipidemia e intolerância à glicose) está relacionado a maior incidência e progressão da OAJ. Sowers et al. (2009) demonstraram prevalência duas vezes maior da OAJ em pacientes com 2 ou mais dos seguintes fatores: HDL \leq 45 mg/dL LDL $>$ 160 mg/dL, triglicérides \geq 200 mg/dL, proteína C-reativa \geq 2 mg/L, razão cintura/quadril \geq 0,81, glicose sérica $>$ 126 mg/dL (ou utilização de medicamentos para diabetes mellitus), hipertensão arterial sistêmica (sistólica $>$ 135 mmHg ou diastólica $>$ 85 mmHg ou uso de medicamentos anti-hipertensivos).

Estudos recentes demonstram associação de fatores genéticos com a incidência e progressão da OA (em uma ou mais articulações) (VAN MEURS, 2017). Algumas mutações e variações genéticas foram associadas à OA precoce ou tardia. Tal contribuição é menos evidente na OAJ, em que predominam os fatores ambientais.

1.2.2.3 Repercussões

A OAJ pode causar dor, rigidez articular, redução na amplitude de movimento e fraqueza muscular (SHARMA, 2021). Os sintomas não se apresentam de forma regular ou previsível, e sim com períodos de intermitência, sendo isso um dos fatores que limitam sua avaliação adequada em estudos longitudinais (SHARMA, 2021). As

consequências no longo-prazo incluem redução no grau de atividade física, alterações do sono, depressão e incapacidade (SHARMA, 2021).

Além dos efeitos na saúde, a OA, especialmente de quadril e joelho, acarreta custos econômicos e sociais de difícil mensuração. Hunter, Schofield e Callander (2014) dividiram os efeitos da OA entre custos diretos, indiretos e intangíveis (Figura 1). Dentre os custos diretos, estão os tratamentos não-farmacológicos e farmacológicos, cirurgias, eventuais efeitos adversos do tratamento e cuidado de saúde de longo prazo. Os custos indiretos incluem absenteísmo, menor empregabilidade e produtividade laboral, aposentadoria precoce, necessidade de cuidados adicionais por outras pessoas (familiares ou colaboradores) e morte prematura. Por fim, os autores destacam os custos intangíveis, que incluem dor, limitação da atividade, redução da qualidade de vida, fadiga e menor participação social.

Enquanto os custos diretos relacionados à OA são mais facilmente calculáveis, é necessário ter atenção também aos custos indiretos e intangíveis, normalmente muito maiores que os diretos (SCHOFIELD et al., 2013). Os custos totais relacionados à OA podem chegar a 1% a 2,5% do produto interno bruto de países desenvolvidos (MARCH; BACHMEIER, 1997).

Estima-se tendência de aumento em tais repercussões (diretas e indiretas), relacionada ao aumento na prevalência de sobrepeso e obesidade e ao envelhecimento da população (HUNTER; SCHOFIELD; CALLANDER, 2014).

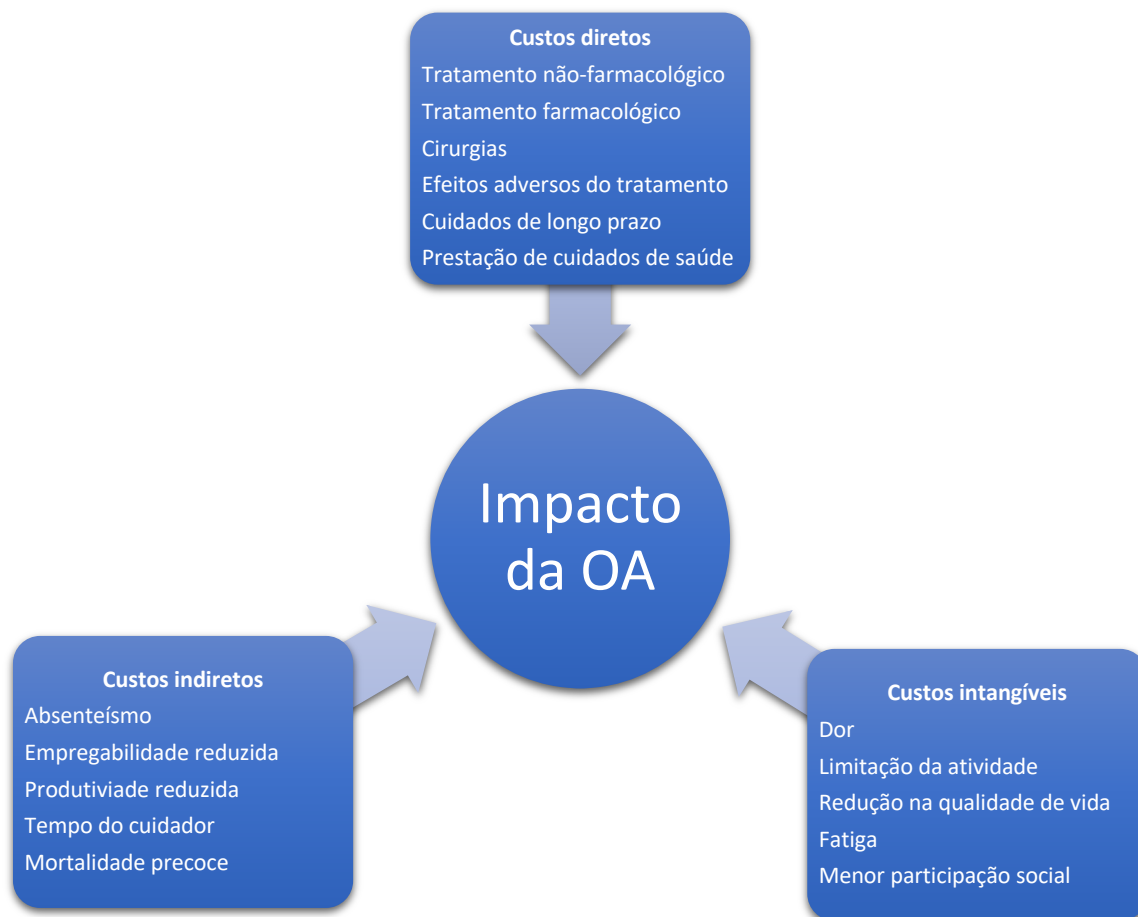


Figura 1: Custos diretos, indiretos e intangíveis relacionados à osteoartrite (traduzido e adaptado de Hunter, Schofield e Callander (2014)).

1.2.2.4 Fisiopatogenia

Anteriormente entendida como uma degeneração articular passiva, a OA é mais recentemente compreendida como uma doença multifatorial de complexa fisiopatogenia, afetando múltiplas articulações e as diversas estruturas que as compõem incluindo, cartilagem, ossos, sinóvia, ligamentos, gordura, meniscos e músculos (LOESER et al., 2012).

Kraus et al. (2015, p. 1237), em consonância com a *Osteoarthritis Research Society International* (OARSI), define osteoartrite como:

“(...) um distúrbio das articulações móveis, caracterizada pelo estresse celular e degradação da matriz extracelular, iniciada por micro e macrotraumas que ativam respostas de reparo mal adaptadas, incluindo mecanismos pró-inflamatórios e do sistema imune inato. A doença se manifesta inicialmente pela desregulação molecular (alterações no metabolismo tecidual articular), seguida de alterações anatômicas e/ou fisiológicas (caracterizadas por degradação da cartilagem articular, remodelação óssea, formação de osteófitos, inflamação articular e perda da função normal da articulação), culminando na doença (tradução nossa).”

Tais processos configuram as fases molecular, pré-radiográfica, e radiográfica da doença (Figura 2).

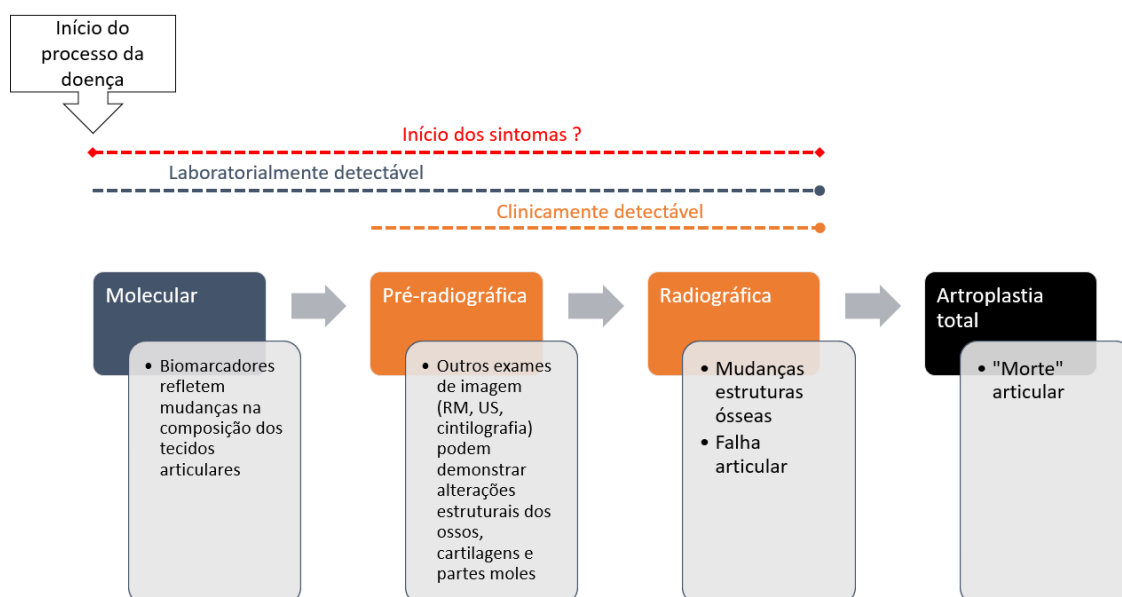


Figura 2: A história natural da osteoartrite e seus biomarcadores (adaptado de Hunter et al. (2013) (apud KRAUS, 2009).

A patogênese complexa da osteoartrite envolve fatores mecânicos, inflamatórios e metabólicos, que resultam em destruição e falha da articulação sinovial. Tais processos são dinâmicos, relacionados a um desequilíbrio entre os mecanismos de reparo e destruição dos tecidos articulares (HUNTER; BIERMA-ZEINSTR, 2019).

Na OA ocorrem mudanças nas propriedades dos materiais que compõem a cartilagem articular, o que aumenta a susceptibilidade a rompimentos por forças físicas (LOESER; COLLINS; DIEKMAN, 2016). Como explicam Hunter e Bierma-Zeinstra (2019), tais erosões ocorrem inicialmente na superfície, aprofundando-se às zonas média e profunda da cartilagem com o passar do tempo, seguindo-se expansão da zona de cartilagem calcificada. Há, então, estímulo à atividade sintetizadora dos condrócitos hipertróficos, numa tentativa de reparar o dano da cartilagem. Tal processo, todavia, estimula mediadores proinflamatórios que desregulam a função do condrócito e ocasionam respostas proliferativas e proinflamatórias da sinóvia adjacente, com predomínio de macrófagos (WOOD et al., 2019). Os sinoviócitos em proliferação também liberam produtos proinflamatórios.

Diversas citocinas proinflamatórias e fatores de crescimento já foram identificados na OA, destacando-se a interleucina 6, *proteína* quimiotática de *monócitos-1*, fator de crescimento endotelial vascular, *proteína* induzida por interferon γ e monoquina induzida por interferon γ (SOHN et al., 2012).

Tais fatores proinflamatórios estimulam a produção de enzimas degradadoras de matriz, incluindo metaloproteinases de matriz (VAN DEN BOSCH; VAN LENT; VAN DER KRAAN, 2020), e sobrecarregam os fatores de crescimento, que não conseguem estimular adequadamente o reparo tecidual. Esse desequilíbrio culmina no predomínio de lesões degenerativas.

Simultaneamente, tem-se hipertrofia tecidual e aumento da vascularização. No osso subcondral ocorre, então, um aumento do *turnover* ósseo, com invasão vascular da cartilagem. Associam-se, ao processo de remodelamento e reparo ósseo, lesões da medula óssea subcondral, com fragmentos ósseos, necrose, fibrose e adipócitos anormais (TALJANOVIC et al., 2008). Osteófitos se desenvolvem nas margens articulares a partir da reativação da ossificação endocondral, sendo afetados não só por fatores inflamatórios biológicos – como o fator de transformação do crescimento β e *proteína* morfogenética óssea 2 – mas também por sobrecarga e alterações da cinética articular.

1.2.2.5 Diagnóstico

De acordo com Sharma (2021), a OAJ pode ser tipicamente diagnosticada com base na anamnese e exame físico. Os sintomas se iniciam gradualmente, destacando-se dor frequente, rigidez articular matinal de curta duração (até 30 minutos) e limitação funcional. O acometimento pode ser uni ou bilateral. A dor pode ser localizada ou difusa, habitualmente apresentando alívio no repouso nas fases iniciais da doença.

Ao exame físico, pode apresentar-se com crepitação, expansão óssea, redução na amplitude de flexão dos joelhos e dor ou sensibilidade à palpação. Derrames articulares estão normalmente ausentes ou de pequeno volume. Katz, Arant e Loeser (2021) destacam, dentre esses sinais, a hipertrofia óssea por sua maior especificidade (95%), apesar de baixa sensibilidade (55%), enquanto crepitação com a movimentação passiva apresenta sensibilidade superior (89%) e especificidade inferior (58%), ainda que significativa. Tais achados podem ser associados entre si e com alterações radiográficas para alcançar maior acurácia para o diagnóstico de OAJ, conforme verificado na Tabela 1.

Tabela 1: Performance dos principais achados do exame físico e radiográficos para o diagnóstico de osteoartrite dos joelhos (adaptado de (KATZ; ARANT; LOESER, 2021)).

Achados	Sensibilidade (%)	Especificidade (%)	Razão de probabilidade (%)
Expansão óssea	55	95	11,0
Crepitação à movimentação passiva	89	58	2,1
Osteófitos	91	83	5,4
Dor articular + osteófitos	83	93	11,9

Como destacam Katz, Arant e Loeser (2021), alterações inflamatórias mais acentuadas, como dor aguda, rigidez de longa duração, derrames articulares de volume moderado a acentuado e aumento da temperatura, apontam para diagnósticos alternativos mais prováveis, como artrites inflamatórias (p. ex.: reumatoide ou psoriática), infecciosas ou por cristais (p. ex.: gota, artrite aguda por pirofosfato de

cálcio). Cabe ainda o diagnóstico diferencial com lesões de partes moles, como bursite, tendinopatia e lesões meniscais.

1.2.2.6 Métodos de imagem

Alterações em exames de imagem podem também contribuir para o diagnóstico da OA. Os achados mais sugestivos incluem osteófitos, esclerose e cistos subcondrais, atrito ósseo e redução de espaço articular, esta relacionada a alterações da cartilagem articular e/ou meniscos. Tais alterações podem ser evidenciadas, em diferentes combinações, nos compartimentos femorotibiais medial, lateral e/ou patelofemoral (SHARMA, 2021).

Radiografia

Segundo Lee et al. (2021), a radiografia é considerada há décadas o método de imagem de escolha para diagnóstico e acompanhamento da OAJ, devido a seu baixo custo e boa capacidade de visualizar as alterações ósseas da doença (ALTMAN; GOLD, 2007; BUCKLAND-WRIGHT, 1994; EMRANI et al., 2008).

A adequada interpretação das radiografias depende de fatores técnicos padronizados (FREDERICK; LANE E; BUCKLAND-WRIGHT, 2002), destacando-se a incidência utilizada, posição dos joelhos em relação ao tubo de raios-X e o filme, angulação e centralização do feixe de raios-X com a articulação e controle da rotação dos joelhos. O controle de tais fatores é necessário a fim de possibilitar a aquisição de radiografias reprodutíveis, em que eventuais alterações visualizadas não decorram da falta de padronização, mas, de fato, de modificações da articulação.

Radiografias dos joelhos são convencionalmente realizadas com o paciente em posição ortostática com os joelhos estendidos (BUCKLAND-WRIGHT, 2006; MAZZUCA; BRANDT; KATZ, 1997). Estudos de biomecânica, todavia, demonstraram que o maior estresse femorotibial ocorre quando o joelho está parcialmente fletido (entre 24 e 28 graus), evidenciando que a radiografia convencional com extensão dos joelhos superestima o espaço articular (MESSIEH; FOWLER; MUNRO, 1990). Isso ocorre pela elevação do fêmur na margem anterior da tíbia. A realização do exame em semiflexão é mais próxima da posição ortostática anatômica, com maior

reprodutibilidade e correlação com lesões da cartilagem articular evidenciadas à artroscopia (MAZZUCA; BRANDT; KATZ, 1997).

A reprodutibilidade na medida do espaço articular depende também da padronização do grau de flexão articular. Diversos autores investigaram a sensibilidade de diferentes graus de flexão dos joelhos na identificação da redução do espaço articular. Resnick e Vint (1980) sugeriram que exames com flexão de 60 a 70 graus apresentam sensibilidade superior àqueles em extensão para a identificação de redução do espaço articular. Messieh, Fowler e Munro (1990) demonstraram a superioridade da avaliação com flexão de 30 graus em relação à extensão. Ravaud et al. (1996) demonstraram que, independentemente do padrão escolhido, pequenas variações (de até 5 graus) no grau de flexão podem ocasionar variações de 10% a 12,5% na estimativa do espaço articular em pessoas sem OA.

Buckland-Wright et al. (1995) sugeriram que a determinação de um grau específico de flexão para cada indivíduo, obtida pela visualização fluoroscópica, poderia aumentar a reprodutibilidade das radiografias seriadas de joelhos. Tal protocolo, todavia, foi considerado pouco custo-efetivo e impraticável para a realidade de estudos longitudinais de grandes populações (NEVITT; FELSON; LESTER, 2017).

Outro fator que deve ser controlado na realização das radiografias dos joelhos é a posição do raio central do feixe de raios-X em relação à articulação (MAZZUCA; BRANDT; KATZ, 1997). Enquanto a imagem é formada sem distorções na tangente raio central, ocorrem distorções em formato de cone das estruturas ao redor desse ponto central, sendo maior quanto maior a distância em relação a esse centro. Mudanças no ângulo do feixe ocasionam, portanto, alterações nas relações e distâncias das margens articulares. O objetivo é radiografar o platô tibial medial o mais horizontalmente e perpendicular aos detectores possível (FREDERICK; LANE E; BUCKLAND-WRIGHT, 2002).

Para controlar essas variáveis e obter exames com adequada acurácia e reprodutibilidade é necessário padronizar a aquisição das radiografias (KOTHARI et al., 2004), especialmente em avaliações seriadas de grandes populações. A OAI (NEVITT; FELSON; LESTER, 2017) propõe a utilização de um protocolo não-fluoroscópico com semiflexão fixa. Esse protocolo apresenta adequada precisão teste-

reteste, comparável a alternativas fluoroscópicas e não-fluoroscópicas (PETERFY et al., 2003).

As incidências mais utilizadas na avaliação dos compartimentos femorotibiais apresentam baixa sensibilidade às alterações do compartimento patelofemoral, sendo indicadas radiografias específicas para essa região (BUCKLAND-WRIGHT, 2006). Dentre elas, destaca-se a incidência lateral em ortostatismo, que permite uma avaliação geral da articulação, bem como da posição vertical da patela, profundidade do sulco troclear, osteofitose marginal e, com menor reprodutibilidade, do espaço articular. Outra incidência que pode acrescentar à avaliação desse compartimento é a axial patelar, usualmente realizada com 20 a 45 graus de flexão, que apresenta melhor delimitação de osteófitos e espaço articular, permitindo ainda a estratificação de OA patelofemoral em lateral e medial (BUCKLAND-WRIGHT, 2006).

O protocolo OAI (NEVITT; FELSON; LESTER, 2017) indica a realização de radiografias em incidência posteroanterior, com flexão fixa dos joelhos de 20 a 30 graus e rotação interna dos pés de 10 graus. O grau de flexão dos joelhos e rotação dos pés deve ser fixado para cada participante com a utilização de um posicionador (por ex.: SynaFlexer™). A avaliação do compartimento patelofemoral deve ser realizada por meio da incidência lateral em posição ortostática.

Diversos sistemas de classificação radiográfica foram propostos. Dentre os mais aceitos e utilizados em estudos científicos, destacam-se as classificações de Kellgren-Lawrence (KL) e OARSI, que incluem a avaliação sistemática de osteófitos e espaço articular, classificando em escala ordinal a gravidade das alterações observadas (ALTMAN; GOLD, 2007; KELLGREN; LAWRENCE, 1957; ODDING et al., 1998).

No atlas de elementos radiográficos individuais da OA, utilizado pela OARSI, Altman e Gold (2007) propuseram a identificação e gradação dos elementos que marcam a OAJ nos compartimentos femorotibiais de forma individual em 0 (normal), 1 (alteração leve), 2 (alteração moderada) e 3 (alteração acentuada). Tal classificação é aplicada a cada um dos seguintes elementos:

- a. osteófitos marginais: nos côndilos femorais medial e lateral e nos platôs tibiais medial e lateral;
- b. redução do espaço articular: nos compartimentos femorotibiais medial e lateral.

Altman e Gold (2007) sugeriram ainda a classificação de outros componentes em ausente ou presente: atrito ou esclerose tibial medial e esclerose femoral lateral.

Kellgren e Lawrence (1957) propuseram um sistema que classifica a OA em 5 graus: (0) Ausência; (1) Duvidosa; (2) Mínima; (3) Moderada; (4) Acentuada. KL grau 1 é definido pela presença de diminutos osteófitos; KL grau 2, pela presença de osteófitos definitivos; KL grau 3, pela presença de osteófitos e moderada redução do espaço articular, associados ou não a deformação óssea; enquanto KL grau 4 é definido pela presença de osteófitos grandes, redução máxima do espaço articular, além de esclerose e deformidade óssea (KATZ; ARANT; LOESER, 2021; YEOH et al., 2021). A figura 3 exemplifica as diferentes classificações KL com radiografias dos joelhos.

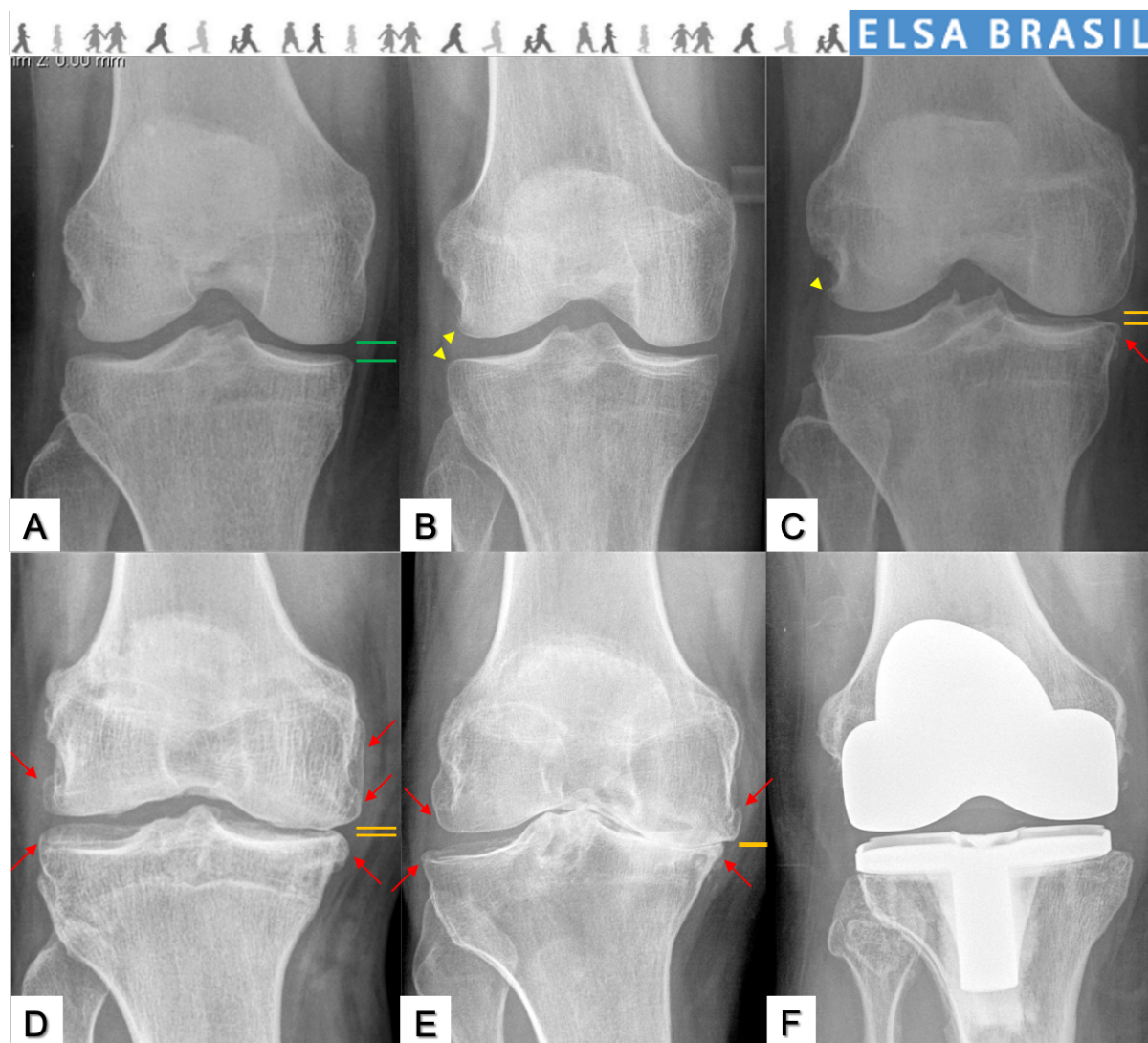


Figura 3: Exemplos de radiografias dos joelhos com diferentes classificações KL.

A: KL 0 – Ausência de osteófitos, espaço articular femorotibial preservado; B: KL 1 – Pequenos osteófitos (cabeças de seta amarelas); C: KL 2 – Osteófito definitivo (seta vermelha) e leve redução do espaço articular (barras laranja); D: KL 3 – Osteófitos definitivos femorais e tibiais, com redução moderada do espaço articular; E: KL 4 – Osteófitos definitivos femorais e tibiais, com redução completa do espaço articular; F: KL não aplicável – Artroplastia total do joelho. *Obs.: redução do espaço articular femorotibial lateral (não exemplificada) também impacta na classificação KL.*

Felson et al. (2011) propuseram algumas modificações para a utilização da classificação KL na identificação de incidência da OAJ (KL grau 2). A OA incidente

seria definida pela presença de osteófitos definitivos associada a redução do espaço articular. O surgimento de osteófito definitivo entre exames seriados sem redução do espaço articular também seria considerado OA incidente, porém com uma nova nomenclatura (KL grau 2/ost).

Apesar de delimitar adequadamente as estruturas ósseas e o espaço articular, existem limitações na avaliação radiográfica dos joelhos. A radiografia não permite a visualização direta adequada de partes moles, incluindo cartilagem, ligamentos, músculos e tendões (LEE et al., 2021). Amin et al. (2005) demonstraram que 42% dos joelhos com perda de cartilagem articular evidente à ressonância magnética (RM) não apresentavam alterações radiográficas. Alterações radiográficas somente são evidentes em estágios intermediários e avançados da OA (GUERMAZI et al., 2012), estimando-se que a redução de espaço articular femorotibial à radiografia somente se apresente após 11% a 13% de perda da cartilagem (JONES et al., 2004).

Outra limitação do estudo radiográfico é a correlação inconsistente com manifestações clínicas (LEE et al., 2021). Sharma (2021) considera que, nos estágios iniciais da osteoartrite, considera que a discordância esteja provavelmente relacionada a baixa sensibilidade da radiografia nessa fase. Nos estágios mais avançados, todavia, a discrepância entre dor e achados de imagem parece estar mais relacionada a outros fatores, destacando-se diferentes fenótipos de OA, adaptações à dor crônica, redução do nível de atividade ou ainda a outros fatores confundidores como status psicossocial e situação econômica (SHAMR e IMAGING KNEE).

Para contornar possíveis fatores confundidores, que influenciam na experiência dolorosa individual, Neogi et al. (2009) avaliaram a relação entre dor articular e achados radiográficos em participantes de duas coortes dos EUA, considerando apenas os participantes que apresentavam níveis de dor discordantes entre os dois joelhos (ou seja, referiam dor em apenas um dos joelhos). Nesse cenário, os autores demonstraram forte associação (tendência com $p < 0,001$) entre marcadores de gravidade radiográficos (maiores graus da classificação KL e dos caracteres individuais – osteófitos e espaço articular) e maior dor articular (considerando frequência, constância e intensidade da dor).

Análise do espaço articular

A avaliação do espaço articular femorotibial é parte relevante das classificações radiográficas da OAJ (KL, OARSI...), sendo fundamental tanto para o diagnóstico da doença quanto para a verificação de progressão em estudos longitudinais. Adiciona-se que a redução do espaço articular é o único desfecho reconhecido para estudos de fase III envolvendo OA nos EUA (HUNTER et al., 2006). A redução do espaço articular femorotibial é considerada um marcador da degeneração cartilaginosa decorrente da presença de OAJ, apresentando importante papel preditivo na história natural da doença, como por exemplo no principal desfecho negativo observável a longo prazo: substituição da articulação acometida por prótese (artroplastia total do joelho) (ALTMAN; GOLD, 2007; FELSON et al., 2008, 2011).

Tal avaliação é atualmente realizada de forma semiquantitativa na maioria dos estudos clínicos, como proposto por Altman e Gold (2007), unitariamente de 0 a 3, sendo (0) espaço articular sem alterações e (3) redução máxima do espaço articular.

Outros métodos de imagem

A ressonância magnética (RM) também pode contribuir na avaliação da OAJ. Dentre as vantagens desse método seccional, destaca-se a melhor avaliação de partes moles, incluindo cartilagem articular, meniscos, ligamentos, tendões e sinóvia. Tais características permitem o estudo de fatores envolvidos na fisiopatogenia da OA, como alterações degenerativas precoces da cartilagem articular, lesões meniscais e ligamentares, formação de osteófitos, edema ósseo, derrame articular e sinovite (GUERMAZI et al., 2012; LEE et al., 2021).

Mesmo com as vantagens descritas, a RM ainda não é o método de imagem mais utilizado para a avaliação da OA. Isso se deve, em parte, devido à disponibilidade limitada, altos custos e longos tempos de aquisição (WANG; OO; LINKLATER, 2018). Menashe et al. (2012) destacam ainda que a sensibilidade da RM não é superior à do diagnóstico clínico e radiográfico, sendo sua acurácia e reprodutibilidade influenciadas também pela falta de padronização nos protocolos de aquisição e interpretação da RM. Segundo Lee et al. (2021), alterações evidenciadas apenas à RM não influenciam diretamente no tratamento escolhido, apresentando pouca relação com o desfecho.

A ultrassonografia (US) é um método de imagem de relativo baixo custo, portátil, móvel e isento de radiação ionizante que permite a visualização da articulação de forma dinâmica (em movimento e em tempo real). Por tais características, sua utilização para avaliação articular e da OA vem crescendo nos últimos anos (OO; BO, 2016). A US permite a visualização da superfície óssea e osteófitos marginais, cartilagem articular e meniscos. Apresenta ainda sensibilidade superior ao exame clínico e comparável à RM na identificação de sinovite e derrame articular (KEEN; WAKEFIELD; CONAGHAN, 2009; NEVALAINEN et al., 2018). A ultrassonografia é o método de escolha para guiar procedimentos diagnósticos e terapêuticos intra e periarticulares percutâneos (KEEN; WAKEFIELD; CONAGHAN, 2009).

Tal método apresenta, todavia, limitações, não apresentando capacidade de identificar alterações do trabeculado e medula óssea, bem como cistos subcondrais (BACKHAUS et al., 2001), apresentando menor acurácia e reprodutibilidade na avaliação dos espaços articulares (PODLIPSKÁ et al., 2016). A ultrassonografia não é rotineiramente recomendada para o diagnóstico e condução da OA, sendo necessários mais estudos para validar e padronizar esse exame no contexto de OA (LEE et al., 2021).

A tomografia computadorizada (TC) é um método seccional que permite a reconstrução tridimensional dos joelhos. Apresenta acurácia superior à radiografia e RM na avaliação das estruturas ósseas articulares (LEE et al., 2021), com custo e tempo de aquisição inferiores à RM (MATHIESSEN et al., 2016). A TC apresenta, todavia, baixa acurácia na avaliação das partes moles, sendo um exame com maior radiação ionizante que a radiografia (LEE et al., 2021), o que limita a sua utilização tanto em estudos quanto na prática clínica.

1.2.2.7 Critérios diagnósticos

Diversos critérios foram propostos para o diagnóstico de OAJ, não havendo consenso sobre qual apresenta melhor acurácia na prática clínica e em estudos epidemiológicos (MIGUEL et al., 2019). Pereira et al. (2011) verificaram as diferenças nas estimativas de prevalência e incidência da OA em joelhos, quadris e mãos de acordo com os critérios utilizados. Dentre os principais utilizados em estudos epidemiológicos, destacam-se os critérios de OA radiográfica e sintomática. Ressaltam-se ainda os

critérios adotados por organizações como o Colégio Americano de Reumatologia (*American College of Rheumatology – ACR*), a Liga Europeia Contra o Reumatismo (*European League Against Rheumatism European – EULAR*) e o Instituto Nacional de Saúde e Excelência em Cuidados do Reino Unido (*National Institute for Health and Care Excellence – NICE*).

OA radiográfica

A OA radiográfica é definida pela identificação de alterações evidentes nas radiografias dos joelhos, seja nos compartimentos femorotibiais, seja no patelofemoral. O sistema de avaliação radiográfica mais aceito e utilizado é o KL (PEREIRA et al., 2011), sendo considerada OA quando a pontuação é maior ou igual a 2 em pelo menos um dos compartimentos do joelho (CHAISSON et al., 2000; FELSON et al., 1995), com gradação de gravidade (KL 2, 3 ou 4).

OA sintomática

Pacientes que apresentam alterações radiográficas e referem algum dos sintomas principais relacionados à OA (dor, desconforto ou rigidez articular) são considerados portadores de OA sintomática (FELSON et al., 1995; MIGUEL et al., 2019). Como ressaltado por Leyland et al. (2018), é necessário que os questionários sigam padronizações na formatação dos enunciados, permitindo melhor delimitação da frequência e temporalidade da dor, com consequente aumento na sensibilidade e especificidade das escalas. Atualmente, admite-se que a presença de sintoma (dor, desconforto, rigidez) frequente (na maioria dos dias de 30 dias) é a pergunta que melhor caracteriza a sintomatologia da OAJ em estudos epidemiológicos (LEYLAND et al., 2018).

ACR

O ACR sugere o diagnóstico de OAJ com base em dois critérios: clínico e clínico-radiográfico. OAJ clínica é caracterizada pela presença de dor em associação com no mínimo três dos seguintes fatores: (1) idade maior ou igual a 50 anos; (2) rigidez articular matinal inferior a 30 minutos; (3) crepitação; (4) sensibilidade à palpação óssea; (5) alargamento ósseo; (6) ausência de aumento da temperatura palpável. A OAJ clínico-radiográfica foi definida pela presença de osteófitos, dor e pelo menos um

dos seguintes itens 1, 2 e 3 supracitados. Existe ainda a classificação clínico-laboratorial de OA, pouco utilizada em estudos epidemiológicos (ALTMAN et al., 1986; ALTMAN, 1991; HUNTER; BIERMA-ZEINSTRA, 2019).

EULAR

A EULAR define OAJ a partir da associação de sintomas (dor persistente, rigidez articular matinal de curta duração e limitação funcional) e sinais (crepitação, restrição ao movimento e alargamento ósseo). Foi demonstrado aumento na probabilidade de OA radiográfica com a associação dos achados clínicos positivos, alcançando 99% na presença de todos os fatores (assumindo prevalência da doença de 12,5% em adultos com idade maior ou igual a 45 anos) (ZHANG et al., 2010).

NICE

O NICE propõe o diagnóstico de OA caso o paciente apresente todos os seguintes critérios: (1) idade igual ou superior a 45 anos, (2) relato de dor articular durante caminhada e/ou subida de degraus e (3) ausência ou duração inferior a 60 minutos da rigidez articular matinal.

Na prática clínica, o diagnóstico da OAJ é eminentemente clínico, baseado na anamnese e exame físico. Hunter e Bierma-Zeinstra (2019) recomendam a padronização do diagnóstico clínico de OAJ com base nas recomendações da ACR e EULAR. Apesar da maior acurácia do critério OA radiográfica, são necessários mais estudos e estabelecimento de novos critérios que permitam a identificação de pacientes com doença em estágio inicial e com maior risco de desenvolvimento de OA (HUNTER; BIERMA-ZEINSTRA, 2019).

1.2.3 Estudos longitudinais sobre Osteoartrite

As principais pesquisas que avaliam a história natural e fatores de risco para o surgimento e progressão da OA são as coortes longitudinais de grande escala, destacando-se o *Framingham Osteoarthritis Study* (FOS), a *Osteoarthritis Initiative* (OAI) e o *Multicenter Osteoarthritis Study* (MOST).

O FOS é um estudo populacional que examinou 1805 participantes, já acompanhados desde 1948 pelo *Framingham Heart Study* (FHS), para sintomas articulares (como dor

e disfunção). Dos participantes que responderam aos questionários, 1424 (79%), com idade entre 63 e 94 anos (média de 73 anos), foram submetidos a radiografia anteroposterior dos joelhos, entre 1983 e 1985 (FELSON, 1990; FELSON et al., 1987). O FOS avalia também participantes da *Framingham Offspring Cohort*, composta por descendentes do FHS, tendo realizado avaliações clínicas, laboratoriais e radiográficas em 1992-1995 (linha de base) e 2002-2005 (NIU et al., 2017). Os dados brutos oriundos do FOS (notadamente as radiografias) não estão disponíveis para consulta nas bases pesquisadas.

A OAI é uma coorte multicêntrica longitudinal observacional com enfoque ao estudo da OAJ iniciada em 2002. Participantes com fatores de risco ou evidências de OA inicial foram recrutados (n=4796) em centros de cinco estados dos EUA (LESTER, 2012). A população do estudo é composta por indivíduos com idades entre 45 e 79 anos no momento do recrutamento, 58% do sexo feminino, incluindo minorias étnicas. Foram realizadas entrevistas e exames físicos, bem como exames de imagem (radiografia e RM) e laboratoriais (sangue e urina), em visitas seriadas (linha de base e 12, 24, 36, 48, 72 e 96 meses) sendo os dados compilados e disponibilizados do repositório online (<https://nda.nih.gov/oai>).

O MOST é um estudo longitudinal observacional prospectivo que avaliou 3026 participantes (60% do sexo feminino), com idade entre 50 e 79 anos, em dois centros (nas cidades de Iowa – IA e Birmingham – AL) (SEGAL et al., 2013). Os participantes selecionados apresentavam diagnóstico prévio de OAJ (um terço) ou alto risco para o desenvolvimento da doença (dois terços), tendo sido submetidos a entrevistas, exames clínicos, exames de imagem e laboratoriais em avaliações seriadas (linha de base e 15, 30, 60, 72 e 80 meses). Os dados eram divulgados no repositório do projeto (<https://most.ucsf.edu/>) porém não estão mais disponíveis com o seu encerramento.

1.2.3.1 Investigação de Distúrbios Musculoesqueléticos no ELSA-Brasil: ELSA-Brasil MSK

A partir de 2012, a avaliação de doenças musculoesqueléticas foi incorporada ao ELSA-Brasil por meio do estudo ancilar ELSA-Brasil MSK. Este acontece no centro de investigação de Minas Gerais, sendo composto por 2901 servidores ativos e aposentados (média de idade de 56,0 anos, 52,9% mulheres). Além das avaliações já

realizadas pelo ELSA-Brasil, o ELSA-Brasil MSK passou a incorporar questionários de incapacidade e sintomas musculoesqueléticos, verificação de fatores de risco para doenças musculoesqueléticas, testes de performance física e radiografias das mãos e joelhos (TELLES et al., 2022). As avaliações do ELSA-Brasil MSK já aconteceram em 2 ondas (2012-2014 e 2016-2018), tendo sido iniciada a 3ª onda de avaliações em agosto de 2022.

Seguindo rigoroso protocolo de coleta de dados do ELSA-Brasil, que garante qualidade das informações, o protocolo de aquisição radiográfica foi avaliado em um estudo anterior com delineamento teste-reteste, tendo sido demonstradas adequada qualidade da imagem e repetibilidade de parâmetros quantitativos (TELLES et al., 2017). Tal protocolo é compatível com os praticados em outros estudos longitudinais como OAI e MOST (NEVITT; FELSON; LESTER, 2017; SEGAL et al., 2013). Os diferentes critérios de classificação da OAJ utilizados no ELSA-Brasil MSK também já foram analisados, com melhor desempenho relacionado ao critério de OA radiográfica, demonstrando sensibilidade de 51,0% e especificidade de 96,7% (MIGUEL et al., 2019).

Telles et al. (2022) relataram prevalência de OAJ radiográfica e sintomática de 18,1% e 7,8%, respectivamente, nos participantes da linha de base do ELSA-Brasil MSK. Dos pacientes com mais de 45 anos, 28% preenchem critérios de NICE para OAJ.

1.2.4 Inteligência artificial

A inteligência artificial (IA) é um ramo da ciência da computação que busca realizar tarefas e resolver problemas que normalmente requereriam habilidades e inteligência humanas por meio de sistemas computacionais (BINVIGNAT et al., 2022; CHARTRAND et al., 2017). É um termo usado de forma ampla, abrangendo diversas técnicas e abordagens para a realização de tais tarefas.

Dentre os campos de estudo da IA, está incluído o de aprendizado de máquina (AM), do inglês *machine learning*, em que são desenvolvidos algoritmos que conseguem aprender padrões e se aperfeiçoar a partir do banco de dados fornecido (BINVIGNAT et al., 2022). Essa área vem passando por uma revolução com o aumento do poder

de processamento computacional e maior volume de dados acumulados, especialmente nos últimos 10 anos (LO VERCIO et al., 2020).

Tal aprendizado pode ocorrer de forma supervisionada ou não supervisionada. No aprendizado de máquina supervisionado, o algoritmo é treinado com dados rotulados – aspectos relevantes estão destacados na amostra de treino. No aprendizado não supervisionado, por sua vez, o programa aprende com dados não rotulados, não tendo acesso a laudos ou desfechos. O AM pode ser ainda classificado em semisupervisionado, quando o treinamento se dá com dados rotulados e não rotulados, e aprendizado de reforço, em que o programa aprende de forma dinâmica com base em punição ou reforço em ciclos de tentativa e erro. Tais algoritmos estão citados na figura 4. Dentre os métodos de aprendizado de máquina, o mais utilizado na medicina são os de aprendizado supervisionado (LO VERCIO et al., 2020).

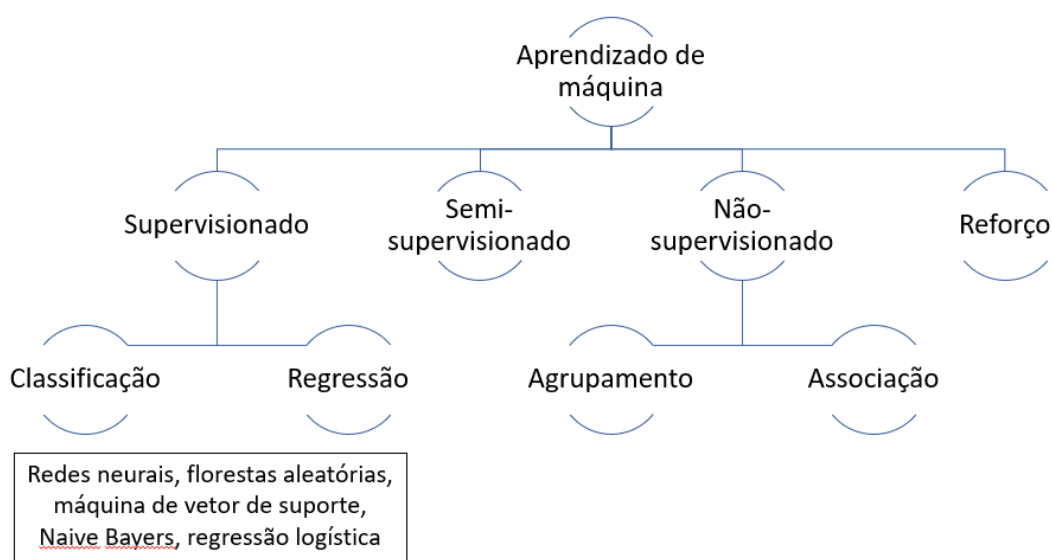


Figura 4: Sumário dos diferentes algoritmos utilizados em aprendizado de máquina (adaptado e traduzido de BINVIGNAT et al., 2022).

Visão computacional é um campo multidisciplinar da IA que estuda e desenvolve sistemas artificiais que permitem a obtenção de informações a partir de imagens. A análise de imagens médicas inclui a segmentação anatômica dos tecidos e estruturas. Esse processo pode ser realizado a partir da delimitação de órgãos (p. ex.: pulmões,

encéfalo, fígado), de porções específicas desses órgãos (p. ex.: lobos pulmonares e cerebrais, segmentos hepáticos), ou ainda de lesões (p. ex.: nódulos, cistos, tumores). Isso pode ser feito de forma semiautomática ou automática (KOENIGKAM SANTOS et al., 2019).

Outro processo envolvido na análise é a extração de atributos, em que características como cor (níveis de cinza), textura e forma são obtidas por meio de algoritmos. Tais programas realizam procedimentos quantitativos que calculam valores numéricos (descritores) para cada característica e as armazenam em um vetor de atributos. A esse processo, segue-se a seleção de atributos relevantes, em que são definidos os atributos mais correlacionados a determinado diagnóstico ou desfecho, geralmente realizada por meio de métodos estatísticos e/ou de aprendizado de máquina (KOENIGKAM SANTOS et al., 2019).

Classicamente, o desenvolvimento de algoritmos de análise de imagens baseava-se em atributos relevantes previamente selecionados. Existe uma abordagem mais recente de AM, todavia, conhecida com aprendizado profundo (AP), do inglês *deep learning*, em que o algoritmo identifica por si só as características que melhor classificariam os dados diretamente das imagens (CHARTRAND et al., 2017). Apesar dessa automatização, técnicas de AP possuem desvantagens como a necessidade de um grande volume de dados para reconhecimento dos padrões, e a dificuldade de se identificar a lógica usada para a classificação (conhecida como *blackbox*) (KOENIGKAM SANTOS et al., 2019).

Existem diversas arquiteturas utilizadas no AP. Dentre aquelas utilizadas na análise de exames de imagem, destacam-se as redes neurais convolucionais (RNC), que têm demonstrado melhores resultados em relação a outras técnicas de IA para essa tarefa, especialmente a partir de 2012 (CHARTRAND et al., 2017). A relação entre áreas da IA está ilustrada na figura 5.

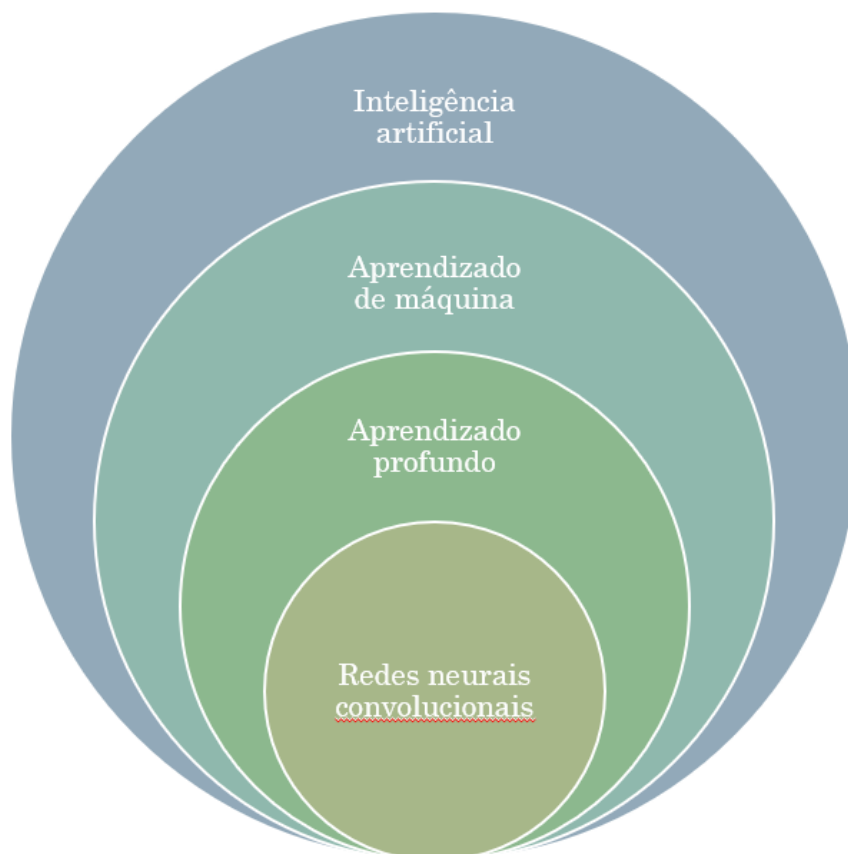


Figura 5: Diagrama de Venn empilhado demonstra a relação entre as áreas da Inteligência Artificial.

Uma RNC (do inglês Convolutional Neural Network – CNN) é um modelo de rede neural artificial (RNA) estruturado sob a forma de camadas interconectadas, cada uma contendo “neurônios artificiais”, conceitualmente análogos aos neurônios biológicos (CHARTRAND et al., 2017). Cada “neurônio artificial” implementa um modelo de classificação que: (1) recebe uma ou mais entradas; (2) pondera essas entradas; (3) submete o resultado a uma função de ativação (p. ex.: linear, sigmoide, gaussiana, etc.), que limita a amplitude dos resultados; e (4) produz um valor de saída. As primeiras camadas extraem e mapeiam o conteúdo da imagem por meio da aplicação de filtros, passando tais atributos para as próximas camadas (KOENIGKAM SANTOS et al., 2019).

Em uma RNC, as primeiras camadas detectam e extraem os atributos primitivos das imagens (tais como bordas e elementos de textura), que são, por sua vez,

processados e selecionados nas camadas seguintes. Tais características são integradas, com diferentes pesos, na camada de saída, que fornece a predição de classe/desfecho de maior probabilidade (KOENIGKAM SANTOS et al., 2019).

A IA vem sendo cada vez mais utilizada para a interpretação de exames de imagem em processos de diagnóstico auxiliado por computador (CAD, do inglês *computer aided diagnosis*), em que o exame é interpretado por um programa que fornece dados que favorecem determinada classificação ou diagnóstico. Doi (2007) define ainda o conceito de diagnóstico computacional automatizado (ACD, do inglês *automated computer diagnosis*), em que o programa fornece o diagnóstico sem a interferência humana.

Diversos estudos têm demonstrado vantagens na aplicação de métodos de aprendizado de máquina, para a classificação e diagnóstico da OA (BINVIGNAT et al., 2022). Tais algoritmos apresentam maior relevância em estudos longitudinais com muitos participantes (BINVIGNAT et al., 2022), como o OAI, MOST e ELSA-Brasil MSK, devido ao grande volume de dados e exames a serem processados e interpretados.

A utilização de RNC na avaliação de radiografias dos joelhos se destaca devido à grande disponibilidade e custo-efetividade desse exame. A radiografia é o principal exame de imagem complementar solicitado na prática clínica para contribuir para o diagnóstico de OA, justificando o enfoque dado a esse método por diversos estudos (YEOH et al., 2021).

Dentre os principais benefícios dessa abordagem, destacam-se a velocidade de leitura e reprodutibilidade. A interpretação de uma radiografia de joelhos, visando classificá-la de acordo com KL, pode demorar ordem de minutos quando realizada por um leitor humano, enquanto leva ordem de segundos ou menos quando feita de forma automatizada. Além disso, o treinamento de pessoas para a leitura de radiografias com boa concordância intra e interobservador (Kappa superior a 0,7) é demorado, podendo levar meses.

Yeoh et al. (2021) avaliaram, em uma revisão sistemática, a utilização de RNC na classificação e segmentação da OA. Dos 12 estudos considerados que aplicam esse método de AP para a classificação de OA em radiografias dos joelhos, apenas 2

utilizavam dados que não os da OAI e MOST: um avaliando a população chinesa (n=1385 exames) e outro a da Coreia do Sul (n=4366 exames) (KIM et al., 2020; LIU; LUO; HUANG, 2020).

Binvignat et al. (2022) publicaram uma revisão sistemática recente sobre a utilização de AM em estudos sobre OA. Os autores demonstram que a maioria dos estudos (85%) estão relacionados à OAJ, com maior enfoque nos exames de imagem (74%) em relação a dados clínicos e biológicos. Quase a metade dos estudos utilizaram os dados da OAI e apenas 7% realizaram validação externa. Nenhum dos estudos incluía a população brasileira ou latino-americana.

De acordo com Binvignat (2022), a aplicação de AM em estudos de OA apresenta algumas limitações. A maioria dos algoritmos utilizados exigem um grande volume de dados rotulados para o seu desenvolvimento. Alguns dos algoritmos de AM disponíveis não permitem o esclarecimento dos fatores que levaram à classificação encontrada (p. ex.: presença de osteófitos, redução do espaço articular, etc.), o que também limita a sua aplicação.

Deve-se conhecer, ainda, o domínio (base de dados) em que a ferramenta foi treinada. As radiografias de estudos longitudinais são geralmente realizadas com posicionamento e incidências específicas, nem sempre utilizadas em contexto não acadêmico, o que limita a acurácia do modelo em exames da prática clínica (BREJNEBØL et al., 2022). A validação externa de ferramentas de IA é fundamental para a sua aplicação prática (BLUEMKE et al., 2020).

2 OBJETIVOS

O presente trabalho pretende avaliar o desempenho de rede neural convolucional previamente publicada (THOMAS et al., 2020) e desenvolver uma outra rede neural convolucional utilizando imagens radiográficas de joelhos dos participantes do ELSA-Brasil MSK.

A RNC proposta por Thomas et al. (2020) foi escolhida por apresentar documentação adequada, que permitiu sua aplicação sobre as radiografias do ELSA-Brasil MSK.

2.1 Objetivos específicos

Artigo 1

- Avaliar a utilização de uma rede neural já publicada e pré-treinada para o diagnóstico de OAJ nas radiografias do ELSA-Brasil MSK;
- Verificar como uma ferramenta de AM desempenha em um domínio diferente do que foi proposto (exames do ELSA-Brasil MSK), que não foram apresentados para a rede na fase de treino.

Artigo 2

- Propor um modelo computacional para classificação de OA em radiografias dos joelhos, desenvolvido apenas com as radiografias do ELSA-Brasil MSK.
- Verificar a capacidade de diagnosticar OA em radiografias dos joelhos de participantes do ELSA-Brasil MSK com diferentes graus de acometimento pela OA;
- Verificar o desempenho da ferramenta (área sob a curva característica de operação do receptor, acurácia, sensibilidade e especificidade) para o diagnóstico de OA;
- Comparar o desempenho do modelo proposto com o de outras redes neurais publicadas com o mesmo objetivo.
- Verificar se o modelo permite triar exames normais com boa sensibilidade.

3 MÉTODOS

3.1 Desenho e Amostra

O presente estudo é uma análise transversal das radiografias da 1ª onda do ELSA-Brasil MSK, realizada concomitantemente à 2ª onda do ELSA-Brasil. Na linha de base, o ELSA-Brasil MSK foi composto por 2901 servidores ativos e aposentados, 53% mulheres, com idades entre 38 e 79 anos (média: 56,0, desvio-padrão: 8,9). Desses, 2830 participantes realizaram radiografias de ambos os joelhos, resultando em imagens de 5660 joelhos (figura 6).

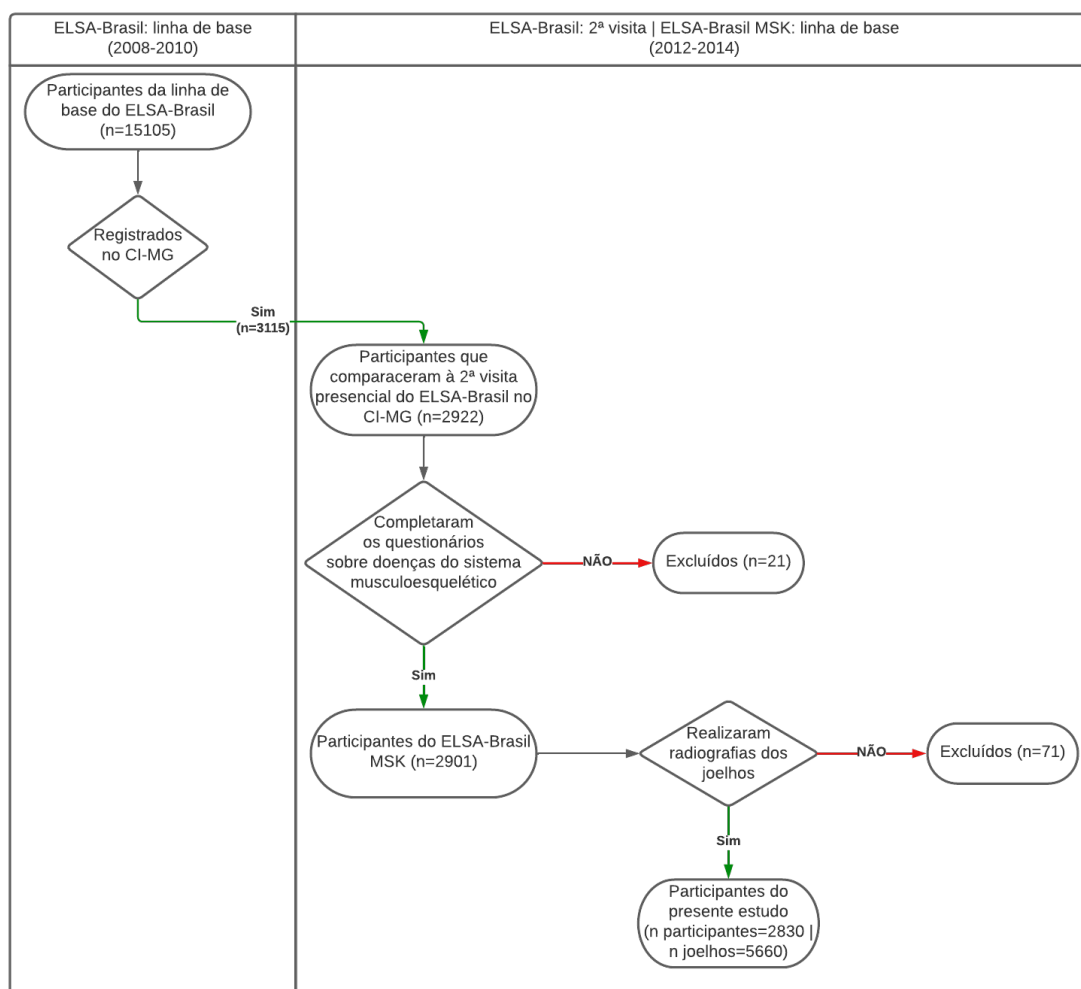


Figura 6: Fluxograma dos participantes incluídos na linha de base do Estudo Longitudinal de Saúde do Adulto – Musculoesquelético (ELSA-Brasil MSK). CI-MG= Centro de Investigação Minas Gerais (adaptado de TELLES et al., 2022).

3.2 Aspectos éticos

Os princípios éticos de respeito à pessoa (assinatura do Termo de Consentimento Livre e Esclarecido - TCLE), beneficência (adequação científica, equilíbrio entre potenciais riscos e benefícios advindos do projeto) e justiça (riscos e benefícios advindos do projeto distribuídos de forma justa) foram considerados no planejamento do ELSA-Brasil e todos os seus estudos ancilares, incluindo o ELSA-Brasil Musculoesquelético.

O ELSA-Brasil possui aprovação da Comissão Nacional de Ética em Pesquisa (CONEP), protocolo 976/2006, e dos Comitês de Ética das seis instituições de ensino e pesquisa participantes. A linha de base do estudo ancilar ELSA-Brasil MSK e o primeiro seguimento deste foram aprovados a partir de emenda submetida ao Comitê de Ética em Pesquisa (CEP) da Universidade Federal de Minas Gerais (UFMG) (emenda CAAE 0186.1.203.000-06 da proposta ETIC 186/06, de 03 de setembro de 2012 e CAAE 47125015.4.1001.5149 parecer 1.897.023 de 24 de janeiro de 2017).

3.3 Exame Radiográfico

As radiografias dos joelhos com processamento digital foram realizadas em uma clínica de Radiologia conveniada ao ELSA-Brasil MSK, localizada nas adjacências do Centro de Investigação ELSA-Brasil Minas Gerais. A aquisição foi realizada em incidência posteroanterior bilateral em semiflexão fixa, com a utilização de posicionador patenteado (INPI BR102013033625-4) desenvolvido pela equipe de pesquisadores do ELSA-Brasil MSK (MACHADO et al., 2013). Todos os exames foram realizados por Técnico ou Tecnólogo em Radiologia devidamente treinado e certificado segundo o protocolo do estudo. O protocolo de aquisição radiográfica foi avaliado em um estudo anterior, tendo sido demonstradas adequada qualidade da imagem e repetibilidade de parâmetros quantitativos (TELLES et al., 2017).

3.4 Interpretação humana

Todos os exames foram interpretados por médico radiologista com treinamento específico para a classificação KL de radiografias de joelho. A concordância da leitura do médico radiologista do ELSA-Brasil MSK em relação a um leitor externo (radiologista musculoesquelético com *background* acadêmico, responsável pelas

leituras do FOS e MOST) e em relação a si próprio foi publicada previamente por Telles et al. (2022), tendo sido estimada em 0,755 (IC 95% 0,663-0,847) e 0,891 (IC 95% 0,807-0,975), respectivamente.

4 RESULTADOS

4.1 Artigo original 1

Title

External validation of a pre-trained Convolutional Neural Network for diagnosis of Osteoarthritis in knee radiographs of the Brazilian Longitudinal Study of Adult Health Musculoskeletal (ELSA-Brasil MSK)

Authors

Júlio G. Domingues¹, Luciana Costa-Silva², Alexei M. C. Machado¹, Luciana A. C. Machado³, Sandhi M. Barreto^{1, 3}, Rosa W. Telles^{1, 3}.

¹ Faculdade de Medicina, Universidade Federal da Minas Gerais, Belo Horizonte, Brazil.

² Instituto Hermes Pardini, Belo Horizonte, Brazil.

³ Hospital das Clínicas-UFMG/EBSERH, Belo Horizonte, Brazil.

Keywords

Osteoarthritis, Knee; Radiography; Neural Networks, Computer; Machine Learning; Diagnosis, Computer-Assisted; Epidemiologic Studies.

Key Points

- This study applies a Convolutional Neural Network (CNN) trained with Osteoarthritis Initiative exams in the Latin American population;
- The CNN had a strong and moderate agreement with radiologists from our cohort study for classification of KL and OA, respectively.
- Model's performance is considered excellent.

Abstract

Osteoarthritis (OA), especially of the knees, is one of the most common diagnoses in general practice, and a leading cause of disability in elderly people. It's usually classified in radiographies with Kallgren-Lawrence (KL) grading system, in a process that can be time-consuming and error-prone, especially considering studies of large populations as ELSA-Brasil, the biggest Brazilian prospective cohort. We analyze the use of a Convolution Neural Network (CNN) trained with Osteoarthritis Initiative (OAI) exams and compare its predictions with ELSA's radiologists in classification of two hundred forty-four knee X-rays obtained between 2012 and 2014. CNN had a strong and moderate agreement with human classification of KL and OA, respectively, with an Area Under the Receiver Operating Characteristics Curve (AUC) of 0.901 for OA diagnosis ($p < 0.001$, CI 95% 0.858-0.945). With some hyperparameter tuning, it can be used in assortment of altered X-rays in ELSA-Brasil MSK.

Introduction

Osteoarthritis (OA) is one of the most common diagnoses in general practice(1) and a leading cause of disability in elderly people(2). Among the many affected sites, the knee OA stands due to its high frequency and associated dysfunction, ahead of other high morbidity causes, such as cardiovascular diseases, diabetes and hip fracture(3).

Knee OA is frequently diagnosed by evaluating radiographies according to the Kellgren-Lawrence (KL) grading system(4), by assessing osteophytes and articular joint-space narrowing, with well established accuracy(5). Nonetheless, the process of reading and classifying radiographies is time-consuming and error-prone, especially considering longitudinal studies of large populations. Therefore, applying an automated machine learning classifier could improve the efficiency of knee OA identification and monitoring over time.

In this study, we evaluated the use of a machine learning model, specifically a Convolutional Neural Network (CNN)(6,7) trained with a dataset of Osteoarthritis Initiative (OAI) exams(8), to identify OA from knee radiographies of the Brazilian Longitudinal Study of Adult Health Musculoskeletal (ELSA-Brasil MSK)(9).

Materials and Methods

Data Collection

ELSA-Brasil MSK is an ancillary study of the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). At baseline (2012–2014), ELSA-Brasil MSK cohort comprised 2901 active or retired civil servants, between 38 and 79 years old (mean age 56.0 years, SD 8.9, 52.9% women), from two teaching and research institutions in the State of Minas Gerais, Brazil(9). Non-fluoroscopic posteroanterior and lateral view fixed-flexion digital bilateral knee radiographies were acquired using computed radiology (ADC-70; Agfa Gevaert NV, Mortsel, Belgium)(10) between 2012 and 2014 using a novel positioning device (INPI BR102013033625-4)(11). The KL grading, ranging from 0 to 4, was used to assess each knee separately. KL grades 0 and 1 are considered negative for OA (OA=0), while 2, 3 and 4 are considered positive (OA=1), with progressively increasing severity(12).

The project was approved by the Ethics and Research Committee (CEP) of Universidade Federal de Minas Gerais (UFMG), Brazil [protocol 186/06].

Reference standard

From ELSA-Brasil MSK's exams collection, a convenience series of 243 films enriched with OA features was selected and evaluated independently by two radiologists (one highly experienced, academically-based bone and joint radiologist - external reader - and another trained and calibrated radiologist responsible for ELSA-Brasil MSK readings), both blind to each other's reading. The interobserver weighted Kappa for the two readers was strong: 0.710 ($p < 0.001$, CI 95% 0.636-0.785) for KL (0&1, 2, 3, 4) e 0.676 ($p < 0.001$, CI 95% 0.583-0.769) for OA (KL \geq 2). In this study, as frequently done in other longitudinal knee OA cohorts, KL 0 and 1 were grouped, as both represent patients without OA.

Radiologists agreed on KL classifications in 190 exams. The remaining 54 exams were then individually evaluated by two other specialists, one radiologist and one rheumatologist, blinded to the other observers, who defined the final diagnosis in a consensual meeting.

Index test

The same radiographies were evaluated by a CNN proposed by Thomas KA et al.(6), with a dense 169-layer architecture, pre-trained with OAI exams (32116 single-limb images). We applied no further tuning nor changed its hyperparameters.

Statistical Analysis

Statistical Analysis were performed with IBM© SPSS© Statistics, version 26.0.0.0. We estimated the weighted Kappa coefficients comparing the results of the CNN with the final diagnosis, separately for KL (0&1, 2, 3 and 4) and OA (absent for KL $<$ 2; present for KL \geq 2) classifications in 243 knees.

Results

The weighted Kappa of the CNN and the final diagnoses, serving as ground-truth, was 0.804 ($p < 0.001$, CI 95% 0.744-0.865) for KL and 0.778 ($p < 0.001$, CI 95% 0.698-0.858) for OA.

Considering KL classification, there was agreement between CNN's classification and the final diagnosis in 206 out of 243 exams (84.8%). Of the 37 discordances:

- in 6 (2.5%) exams, CNN misclassified KL 2 knees as KL 1;
- in 20 exams (8.2%), CNN misclassified KL 0 or 1 knees as KL 2;
- in 9 exams (3.7%), CNN misclassified KL 2 knees as KL 3;
- in 2 exams (0.8%), CNN misclassified KL 4 knees as KL 3.

It is worth noting that all exams classified as KL 0 by the CNN (101 knees) actually presented no OA knees (KL 0 or 1) according to ground-truth. The results are detailed in the confusion matrix below (Table 1).

Table 1: Confusion matrix between CNN's predicted KL classifications related to final diagnosis (ground-truth). CNN = Convolutional Neural Network.

KL		Predicted (CNN)				Total
		0 / 1	2	3	4	
Ground-truth	0 / 1	133 (86.9%)	20 (13.1%)	0	0	153 (100.0%)
	2	6 (10.3%)	43 (74.1%)	9 (15.5%)	0	58 (100.0%)
	3	0	0	26 (100.0%)	0	26 (100.0%)
	4	0	0	2 (33.3%)	4 (66.7%)	6 (100.0%)

Considering OA classification, there was agreement between CNN's classification and the ground-truth in 217 out of 243 exams (89.3%). Six exams (2.5%) were considered false negatives, while twenty (8.2%) were considered false positives (Table 2). CNN's overall sensitivity and specificity for OA diagnosis were, respectively, 0.933 and 0.869, with Area Under the Receiver Operating Characteristics (AUC) for OA diagnosis of 0.901 ($p < 0.001$, CI 95% 0.858-0.945). The positive and negative predictive values were, respectively, 0.808 and 0.957.

Table 2: Confusion matrix between CNN's OA predictions and ground-truth. CNN = convolutional neural network, OA = Osteoarthritis.

OA		Predicted (CNN)		Total
		0	1	
Ground-truth	0	133 (86.9%)	20 (13.1%)	153 (100.0%)
	1	6 (6.67%)	84 (93.3%)	90 (100.0%)

Qualitative results

Figure 1 presents three X-rays that have definite femoral osteophytes (red arrows) according to all human readers, which make them KL 2, although CNN classified them as KL 1 (no OA). Figure 2 presents a case of CNN false positive.

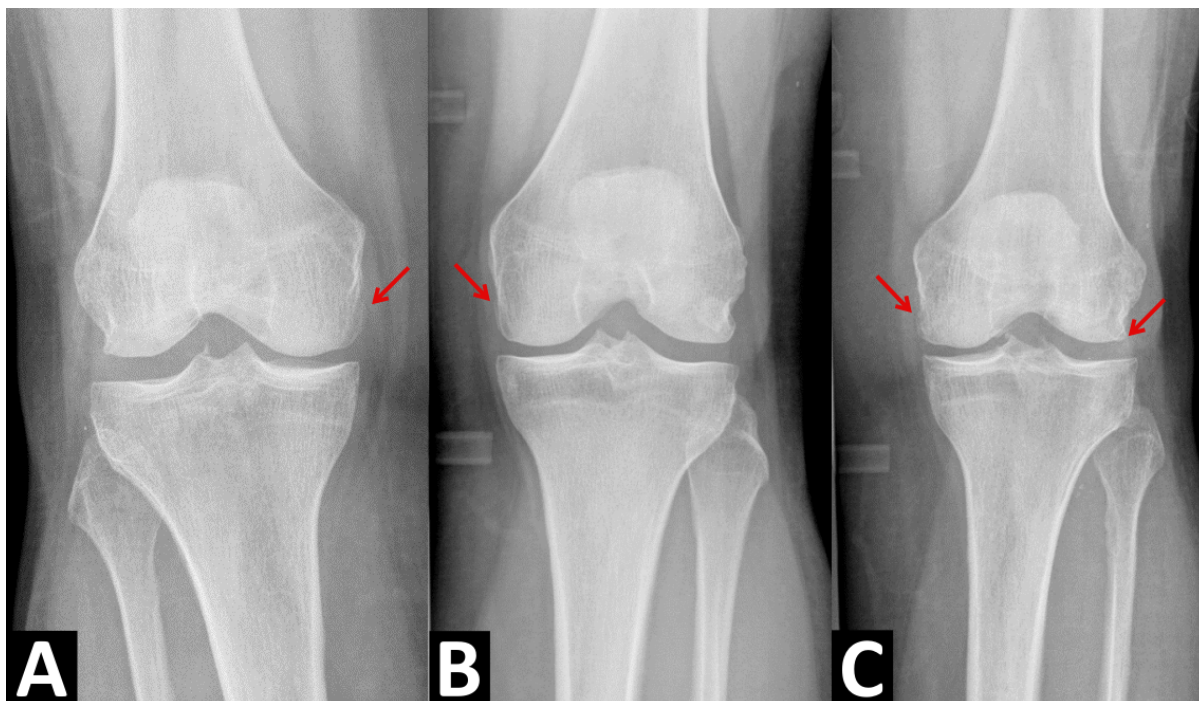


Figure 1: False negatives examples (three different subjects: A, B and C). Femoral definite osteophytes (red arrows) were not considered in CNN's reading. CNN = Convolutional Neural Network.



Figure 2: False positive example: both radiologists considered this a normal exam (KL 0), while CNN classified it as KL 2. CNN = Convolutional Neural Network, KL = Kellgren-Lawrance classification.

Discussion

In longitudinal studies with a large number of participants, systematic human assessment of imaging exams can be a limiting factor, either due to the long time spent in interpreting the images or due to researchers' turnover. The use of automated computer diagnosis, in this context, can contribute to greater flexibility and reproducibility in the interpretation of tests(6). Therefore, it can be used as a screening method, representing a significant saving of resources. It should be noted, however, that neural networks, especially without adequate hyperparameter adjustment, can produce systematic errors, which can result in misclassifications.

In this study, the CNN presented strong and moderate agreement(13), respectively in the classification of KL (0&1, 2, 3 and 4) and OA (KL \geq 2) when compared to human evaluation. The results are promising, especially considering that no changes were

made in the classifier's parameters and that ELSA-Brasil MSK participants are from a country with a different epidemiological profile. The performance of the neural network is considered excellent(14), with accuracy comparable to Thomas K et al. analysis(6) (0.901 vs 0.872, respectively). To our knowledge, this is the first study to apply and validate an OAI trained CNN in the Latin American population.

Considering the applicability of this CNN in the ELSA-Brasil MSK cohort to read X-rays from now on, having an even greater negative predictive value would be preferred. Currently, all the exams are classified by radiologists, thus the use of a software to recognize normal exams would save significant time and resources. Changes in hyperparameters to prioritize sensitivity should improve the algorithm's performance in this scenario, even though this would increase the chance of type 1 error. As a result, very likely normal images would be skipped and those with alterations minimally compatible with OA would be interpreted by the radiologists. Retraining the CNN on total ELSA-Brasil MSK's exams, despite laborious, would likely increase its accuracy. Although the dataset is smaller – about six thousand knees – than the usual for this purpose, the use of data augmentation techniques could mitigate this issue.

While algorithms evaluating imaging exams are in vast expansion today, most studies do not allow for external validation(15). The delivered software often does not have open-source code or sufficient documentation for its application in other institutions, harming the pursuit of translational medicine. It is also noteworthy that many diseases cannot be completely diagnosed or ruled out based on just one imaging exam, and it is often necessary to associate them with clinical data, radiological follow-up or the use of complementary imaging methods.

Conclusions

This study evaluated the use of a pre-trained CNN in radiographies from the Brazilian largest longitudinal epidemiological study of adults. The software had moderate to strong agreement with ELSA's radiologists and accuracy consistent with previously published work. Applying modifications, with a special focus on increasing negative predictive value, should allow its use for assortment of altered X-rays for human

interpretation, speeding up the process of radiography reading and minimizing spurious errors.

The use of pre-trained machine-learning models also foments a discussion on their ability to adapt to other domains(16). In this study, a CNN trained from a set of OAI images was applied, with no adjustments, to a distinct epidemiological population's dataset. As CNNs are considered blackbox methods, such degree of compatibility is highly desired and suggests the feasibility of developing methods that are robust to data and context diversity.

CNN's consistency and coherence are still to be tested in serial exams. The capacity to identify incident or progressive OA might potentialize its use in longitudinal studies. Participants of ELSA-Brasil MSK have attended to two visits, until now, and the third is under way, which will allow the validation of this CNN in such context.

Acknowledgments

To Prof. P. Aliabadi., one of the readers of the images used in this study.

References

1. Santo L, Okeyode T. National Ambulatory Medical Care Survey: 2018 National Summary Tables [Internet]. 2018 [cited 2022 Mar 3]. Available from: https://www.cdc.gov/nchs/data/ahcd/namcs_summary/2018-namcs-web-tables-508.pdf
2. Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PWF, et al. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. *Am J Public Health.* 1994;84(3):351.
3. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet.* 2012;380:2163–96.
4. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis.* 1957;16(4):494–502.
5. Damen J, Schiphof D, Wolde S ten, Cats HA, Bierma-Zeinstra SMA, Oei EHG. Inter-observer reliability for radiographic assessment of early osteoarthritis features: the CHECK (cohort hip and cohort knee) study. *Osteoarthritis Cartilage.* 2014;22(7):969–74.
6. Thomas KA, Kidziński Ł, Halilaj E, Fleming SL, Venkataraman GR, Oei EHG, et al. Automated Classification of Radiographic Knee Osteoarthritis Severity Using Deep Neural Networks. *Radiol Artif Intell.* 2020 Mar 1;2(2):e190065.
7. Indolia S, Goswami AK, Mishra SP, Asopa P. Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach. *Procedia Comput Sci.* 2018 Jan 1;132:679–88.
8. Lester G. The Osteoarthritis Initiative: A NIH Public–Private Partnership. *HSS Journal.* 2012 Feb 23;8(1):62.

9. Telles RW, Machado LAC, Costa-Silva L, Barreto SM. Cohort Profile Update: The Brazilian Longitudinal Study of Adult Health Musculoskeletal (ELSA-Brasil MSK) cohort. *Int J Epidemiol*. 2022 Apr 26;dyac084
10. Telles RW, Costa-Silva L, Machado LAC, Reis RCP dos, Barreto SM. Fixed-flexion knee radiography using a new positioning device produced highly repeatable measurements of joint space width: ELSA-Brasil Musculoskeletal Study (ELSA-Brasil MSK). *Rev Bras Reumatol*. 2017;57(2):154–61.
11. Machado LAC, Costa-Silva L, Barreto SM, Telles RW, Naves EB, Carneiro AMFC. Positioner for the acquisition and quality control of radiographic images of knees in the fixed flexion position (radiolfixer). Brazil: National Institute of Industrial Property; BR102013033625-4, 2013. p. 161–161.
12. Felson DT, Niu J, Guermazi A, Sack B, Aliabadi P. Defining radiographic incidence and progression of knee osteoarthritis: suggested modifications of the Kellgren and Lawrence scale. *Ann Rheum Dis*. 2011 Nov;70(11):1884–6.
13. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276.
14. Carter J v., Pan J, Rai SN, Galandiuk S. ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. *Surgery*. 2016 Jun 1;159(6):1638–45.
15. Binvignat M, Pedoia V, Butte AJ, Louati K, Klatzmann D, Berenbaum F, et al. Use of machine learning in osteoarthritis research: a systematic literature review. *RMD Open*. 2022 Mar 1;8(1):e001998.
16. Oliveira HN, Ferreira E, Santos JAD. Truly generalizable radiograph segmentation with conditional domain adaptation. *IEEE Access*. 2020;8:84037–62.

4.2 Artigo original 2

Título

Desenvolvimento de rede neural convolucional para o diagnóstico radiográfico de osteoartrite dos joelhos no Estudo Longitudinal da Saúde do Adulto – Musculoesquelético (ELSA-Brasil MSK)

Autores

Júlio Guerra Domingues¹, Daniella Castro Araújo^{2, 3}, Luciana Costa-Silva⁴, Alexei Manso Corrêa Machado¹, Luciana Andrade Carneiro Machado⁵, Adriano Alonso Veloso², Sandhi Maria Barreto^{1, 5}, Rosa Weiss Telles^{1, 5}.

¹ Faculdade de Medicina, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte - MG, Brasil

² Departamento de Ciência da Computação, UFMG, Belo Horizonte - MG, Brasil

³ Huna-AI, São Paulo - SP, Brasil

⁴ Instituto Hermes Pardini, Belo Horizonte - MG, Brasil

⁵ Hospital das Clínicas – UFMG / EBSE RH, Belo Horizonte - MG, Brasil

Unitermos

Osteoartrite do Joelho; Radiografia; Redes Neurais de Computação; Aprendizado de Máquina; Diagnóstico por Computador; Estudos Epidemiológicos.

Resumo

Introdução: A osteoartrite dos joelhos (OAJ) é frequentemente diagnosticada por meio da avaliação de radiografias. Em estudos longitudinais de larga escala, esse processo é laborioso e demorado, exigindo treinamento específico. Nesse contexto, a aplicação de modelos computacionais, como as redes neurais convolucionais (RNC), para classificação automática da OAJ pode contribuir para o fluxo de diagnóstico.

Objetivo: Desenvolver uma RNC treinada a partir dos exames da linha de base do ELSA-Brasil MSK. **Método:** Trata-se de um estudo transversal realizado com base nas radiografias da linha de base do ELSA-Brasil MSK (5660 radiografias dos joelhos em incidência posteroanterior). **Resultados:** A RNC desenvolvida apresentou área sob a curva característica de operação do receptor (AUC) de 0,866 (IC 95% 0,842-0,882), com acurácia balanceada máxima de 0,814. O modelo pode ser calibrado para alcançar, não simultaneamente, valores máximos de 0,907 para acurácia, 0,938 para sensibilidade e 0,994 para especificidade. **Conclusões:** A RNC desenvolvida pode ser utilizada como ferramenta de triagem, reduzindo o número total de exames avaliados pelos radiologistas do estudo, e/ou como ferramenta de segunda leitura, contribuindo com a redução de possíveis erros de interpretação.

Introdução

A osteoartrite (OA) é um dos problemas de saúde mais prevalentes na população mundial (1), especialmente na população idosa (2). O acometimento dos joelhos destaca-se não apenas pela alta prevalência, mas também pela morbidade associada, sendo uma das principais causas de anos vividos com incapacidade no mundo (3). Estima-se a incidência da osteoartrite dos joelhos (OAJ) entre 80 e 240/100.000 pessoas-ano nos EUA, com aumento significativo a partir dos 50 anos de idade (4,5). Identificou-se OAJ radiográfica em 18,1% dos participantes do maior estudo longitudinal para estudo de doenças musculoesqueléticas do Brasil (6).

A osteoartrite dos joelhos (OAJ) pode causar dor, rigidez articular, redução na amplitude de movimento e fraqueza muscular (7). As consequências no longo-prazo incluem redução no grau de atividade física, alterações do sono, depressão e incapacidade (7). Cabe ressaltar ainda as repercussões econômicas e sociais relacionadas à OA, incluindo todas as articulações acometidas, que incluem custos diretos (como tratamentos e cirurgias), indiretos (absenteísmo, menor empregabilidade e aposentadoria precoce) e intangíveis (dor, redução da qualidade de vida e menor participação social) (8). March e Bachmeier (9) estimaram que os custos totais relacionados à OA podem chegar a 1% a 2,5% do produto interno bruto de países desenvolvidos (9) e estima-se tendência de aumento em tais custos devido ao aumento na prevalência de sobrepeso e obesidade e ao envelhecimento da população (10).

O diagnóstico da OAJ pode ser realizado a partir de critérios clínicos e radiográficos, em associação ou não, sendo o critério radiográfico considerado dos mais sensíveis (11,12). Em estudos epidemiológicos longitudinais, o diagnóstico é usualmente realizado a partir de achados das radiografias de joelhos (13), sendo o sistema de gradação mais utilizado o de Kellgren e Lawrence (KL) (14). Joelhos com KL 0 e 1 indicam ausência de OAJ definitiva, enquanto escores KL 2, 3 e 4 indicam a presença de OA.

A classificação das radiografias em estudos longitudinais é usualmente realizada por médicos especialistas, e requer um processo de treinamento, padronização e calibração rigoroso (6). A análise das imagens consiste na gradação semiquantitativa

dos osteófitos e espaços articulares, de acordo com atlas radiográfico. Em pesquisas de larga escala, tal processo se torna demasiadamente demorado e oneroso, estando sujeito ao grau de experiência dos observadores. Por esta razão, estudos têm sido desenvolvidos com o objetivo de verificar a viabilidade na utilização de modelos computacionais para classificação automática e semiautomática da OA de joelhos (15), reduzindo o número total de exames a serem avaliados por humanos (16).

Dentre os modelos de classificação automática, destacam-se as redes neurais convolucionais (RNC), que têm demonstrado resultados promissores na área de diagnóstico médico por imagem (16). A utilização de tais redes permitiria maior rapidez e reprodutibilidade das leituras (16). O treinamento e verificação da acurácia desses modelos computacionais, todavia, concentram-se em estudos clínico-epidemiológicos dos Estados Unidos (15,17–24), sendo escassas as ferramentas validadas na população de outros países. Duas revisões recentes sobre o tema (15,17) não identificaram estudos que abrangessem a população brasileira, ou mesmo latino-americana, utilizados no treinamento das RNC atualmente existentes, tendo demonstrado a necessidade maior validação externa.

O Estudo Longitudinal de Saúde do Adulto (ELSA-Brasil), maior estudo epidemiológico longitudinal da América Latina (25), abrange, desde 2012, a avaliação de doenças musculoesqueléticas, por meio do estudo ancilar ELSA-Brasil Musculoesquelético (ELSA-Brasil MSK). Este acontece no centro de investigação de Minas Gerais (CI-MG), e incorpora, além das avaliações já realizadas pelo ELSA-Brasil, questionários de incapacidade e sintomas musculoesqueléticos, verificação de fatores de risco para doenças musculoesqueléticas, testes de performance física e radiografias das mãos e joelhos (6).

O objetivo desse trabalho é propor um modelo computacional para classificação de OA em radiografias de joelho, treinado com as radiografias do ELSA-Brasil MSK.

Materiais e métodos

Amostra

O desenvolvimento do modelo computacional para análise automática de radiografias foi realizado a partir dos exames realizados na primeira visita (linha de base) da coorte ELSA-Brasil MSK. O ELSA-Brasil MSK incluiu, em sua linha de base, 2901 servidores avaliados no CI-MG, ativos ou aposentados de duas grandes instituições de ensino e pesquisa, com idades entre 38 e 79 anos (média de 56,0 anos), sendo 52,9% mulheres. Desses participantes, 2830 realizaram radiografias de ambos os joelhos, resultando em imagens de 5660 joelhos. Detalhes sobre o delineamento e perfil da coorte ELSA-Brasil MSK foram publicados previamente (6).

Exame Radiográfico

As radiografias dos joelhos com processamento digital foram realizadas em uma clínica de Radiologia conveniada ao ELSA-Brasil, localizada nas adjacências do CI-MG. A aquisição foi realizada em incidência posteroanterior bilateral em semiflexão fixa, com a utilização de posicionador (INPI BR102013033625-4) desenvolvido pela equipe de pesquisadores do ELSA-Brasil MSK (26). Todos os exames foram realizados por técnico ou tecnólogo em radiologia devidamente treinado e certificado segundo o protocolo do estudo.

O protocolo de aquisição radiográfica foi avaliado em um estudo anterior com delineamento teste-reteste, tendo sido demonstradas adequada qualidade da imagem e repetibilidade de parâmetros quantitativos (27). Tal protocolo é compatível com os praticados em outros estudos longitudinais como *Osteoarthritis Initiative* (OAI) e *Multicenter Osteoarthritis Study* (MOST) (28,29).

Interpretação humana

A interpretação das radiografias foi realizada de acordo com o seguinte protocolo, previamente publicado e validado (6): (i) triagem de “OA possível” por dois tecnólogos de forma independente; e (ii) todos os exames categorizados, por pelo menos um dos tecnólogos, como “OA possível” foram revisados por médico radiologista com treinamento específico. A concordância da leitura do médico radiologista do ELSA-Brasil MSK em relação a um leitor externo (radiologista musculoesquelético com

background acadêmico, responsável pelas leituras do *Framingham Osteoarthritis Study* - FOS e MOST) e em relação a si próprio foi publicada previamente por Telles et al. (6), tendo sido estimado kappa interobservador de 0,755 (IC 95% 0,663-0,847) e intraobservador de 0,891 (IC 95% 0,807-0,975) para diagnóstico radiográfico da OAJ.

Radiografias KL 0 e 1 foram consideradas negativas para OA; radiografias KL 2, 3 e 4 foram consideradas positivas para OA. A classificação binária (OA=0; OA=1) foi utilizada como valor de referência pela rede neural.

Modelo Computacional

Avaliando-se as técnicas de IA mais utilizadas para a avaliação de imagens médicas atualmente (16,30,31), destaca-se a arquitetura convolucional densamente conectada (DenseNet). O modelo de RNC aqui proposto utiliza arquitetura DenseNet pré-treinada, de 161 camadas (figura 1).

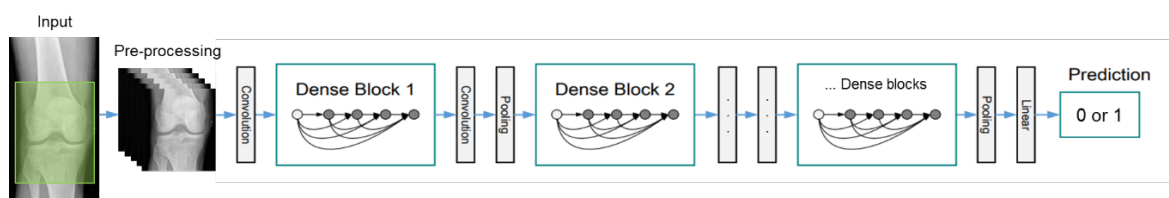


Figura 1: Ilustração esquemática da arquitetura DenseNet. Os pares de camadas são conectados, permitindo que elementos das primeiras camadas (como bordas) possam ser também utilizados nas últimas (adaptado de Huang et al. (32)).

Foi realizado pré-processamento das imagens a partir dos dados brutos das radiografias em PA bilateral dos joelhos (arquivos DICOM). Inicialmente foram individualizados os joelhos direito e esquerdo, seguida de ampliação e redimensionamento das imagens, em matriz quadrada, com localização das regiões de interesse (compartimentos femorotibiais) (Figura 2).

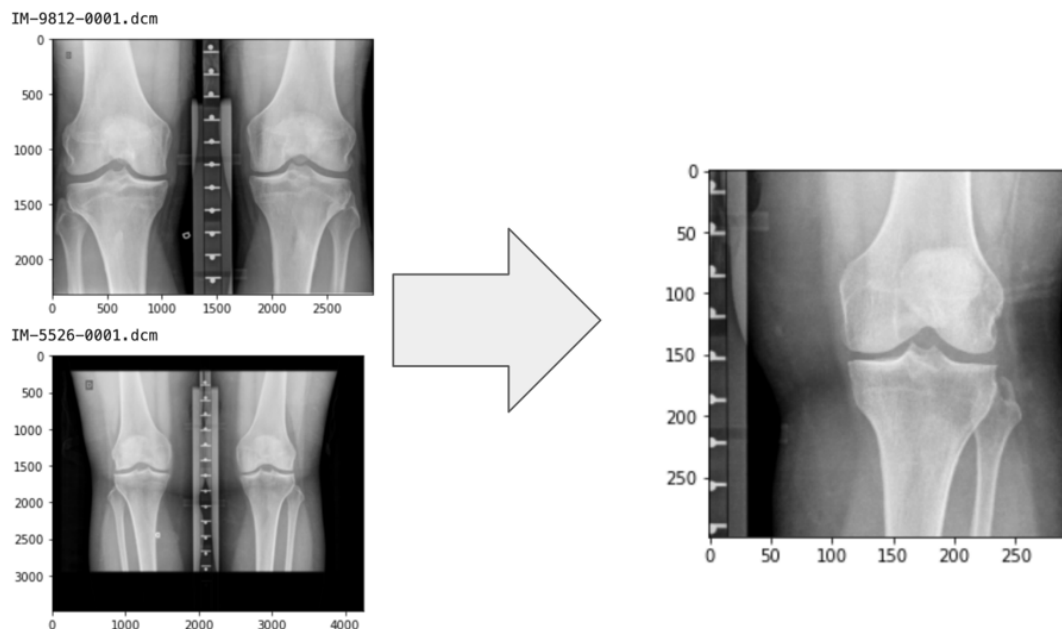


Figura 2: Demonstração do pré-processamento

Para aumento do número de imagens disponíveis para o treinamento da rede neural, foram realizados os seguintes mecanismos de *data augmentation* randômico do módulo `torchvision.transforms` da biblioteca PyTorch, aplicados sobre a amostra de treinamento: rotação ($0,5^\circ$) e desfoque gaussiano; inversão horizontal, ajuste de nitidez (fator=0,5) e desfoque gaussiano. Seguiram-se redimensionamento, corte central (CenterCrop), além de normalização. Os exames da amostra foram divididos em duas partições mutuamente exclusivas (*folds*) de treino e teste.

Como a saída do modelo é uma probabilidade para cada imagem, ele pode ser calibrado por meio da otimização de limiares (*threshold-moving*), que variam de 0 a 1. Em problemas de classificação binária, o limiar de decisão padrão é de 0,5: caso a probabilidade seja superior a esse valor, considera-se a classe 1, e caso contrário a classe 0.

Análise dos dados

As classificações binárias (OA=0 | OA=1) realizadas pela rede neural foram comparadas com a análise do radiologista (valor de referência). Realizou-se a verificação do desempenho da RNC utilizando o módulo *metrics* da biblioteca *scikitlearn* (versão 1.0.2). Para cada limiar, as proporções de resultados verdadeiro positivos, verdadeiro negativos, falso positivos e falso negativos foram armazenadas em vetores, a partir dos quais foram calculados os valores de sensibilidade, especificidade, precisão, acurácia, acurácia balanceada, acurácia balanceada ponderada e escores F1 e F2, considerando-se a média das *folds*.

A acurácia é calculada pela razão entre o número de acertos (verdadeiro positivos e verdadeiro negativos) e o total de exames avaliados. Em amostras desbalanceadas, todavia, como no caso do presente estudo, em que existem muito mais exemplos de exames normais que alterados, essa métrica pode não demonstrar adequadamente o desempenho do modelo. Nesse contexto, a utilização da acurácia balanceada (AB) permite uma melhor estimativa do rendimento da RNC (33), sendo calculada segundo a fórmula: $\frac{\text{sensibilidade} + \text{especificidade}}{2}$. Alguns autores defendem, ainda, a utilização da acurácia balanceada ponderada (ABP, do inglês *Weighted Balanced Accuracy*) (33,34), que permite a atribuição de diferentes pesos para cada métrica, tendo sido calculada da seguinte forma: $\frac{(2 * \text{sensibilidade}) + \text{especificidade}}{3}$.

Para o cálculo da área sob a curva das características de operação do receptor (AUC) do modelo foram consideradas as probabilidades preditas para cada imagem, calculadas e armazenados em lotes de 128 exames, que foram comparados com o valor-verdade por meio da função *roc_auc_score* do scikit-learn. Essa função traça a taxa de positivos classificados corretamente entre todas as previsões positivas (ou seja, a taxa de verdadeiros positivos) em função de positivos incorretos entre todos os negativos (ou seja, a taxa de falsos positivos), em limites variados (35).

Resultados

A rede neural desenvolvida apresentou, considerando a média simples dos dois *folds*, acurácia de 0,814 (no ponto de acurácia balanceada máxima), com sensibilidade de 0,755 e especificidade de 0,821. A AUC do modelo foi de 0,866 (IC 95% 0,854-0,883) (figura 3).

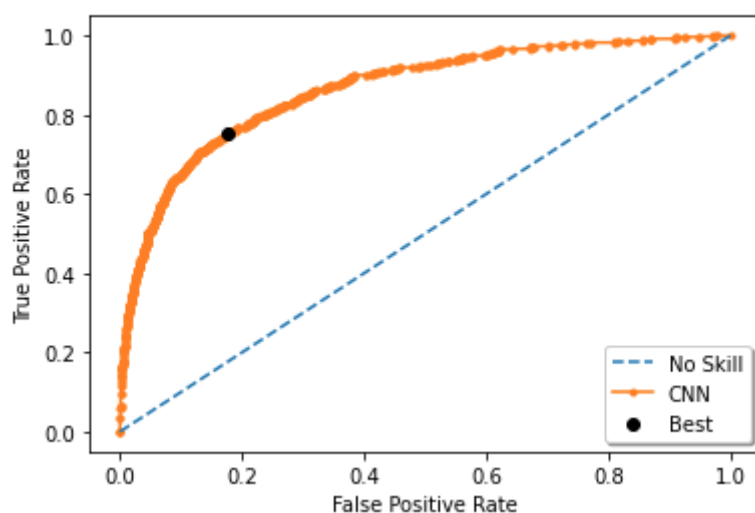


Figura 3: Curva característica de operação do receptor do modelo.

Seguindo a técnica acima exposta, o modelo pode ser calibrado para alcançar, não simultaneamente, valores máximos de 0,907 para acurácia, 0,938 para sensibilidade e 0,994 para especificidade. Os valores máximos dos escores F1 e F2 alcançados foram de, respectivamente, 0,553 e 0,619. A tabela 1 demonstra os valores máximos alcançados pela RNC, de acordo com a métrica otimizada.

Tabela 1: Valores de acurácia, acurácia balanceada, acurácia balanceada ponderada, sensibilidade, especificidade, precisão e escores F1 e F2 para cada limiar definido. Acu = Acurácia, AB = Acurácia balanceada, ABP = Acurácia balanceada ponderada, Sens = Sensibilidade, Esp = Especificidade, Prec = Precisão, F1 = escore F1, F2 = escore F2.

Métrica maximizada	Limiar	Acu	AB	ABP	Sens	Esp	Prec	F1	F2
Sensibilidade	0,010	0,400	0,646	0,755	0,973	0,319	0,758	0,272	0,479
ABP	0,040	0,649	0,758	0,805	0,899	0,617	0,235	0,373	0,575
AB	0,140	0,814	0,789	0,778	0,756	0,822	0,357	0,485	0,618
F1	0,314	0,881	0,774	0,728	0,636	0,913	0,490	0,553	0,600
Acurácia	0,712	0,907	0,671	0,568	0,363	0,978	0,687	0,475	0,401
Especificidade	0,900	0,902	0,599	0,467	0,203	0,994	0,806	0,325	0,239

Discussão

O modelo desenvolvido apresentou bom desempenho (36) para o diagnóstico de OA em radiografias de joelhos. A definição do limiar, com conseqüente priorização de alguma das métricas de desempenho, deve ser realizada de acordo com o contexto em que será utilizada.

A comparação entre a eficiência de diferentes modelos de inteligência artificial não está ainda padronizada na literatura (37). Apesar de haver uma predileção histórica pela acurácia, considera-se a AUC a métrica mais adequada atualmente para avaliação de desempenho (37,38).

Na principal revisão recente sobre a utilização de algoritmos de aprendizado de máquina para a avaliação de OA, Binignat et al. (15) descrevem apenas dois estudos que se propuseram a diagnosticar OAJ somente em radiografias (24). Brahim et al. (24) alcançaram 82.98% de acurácia (com 87,15% de sensibilidade e 80,65% de especificidade) para a diferenciação de KL 0 e 2, tendo sido treinado em 1024 imagens (514 KL=0; 514 KL=2) da OAI. Não foi calculada a AUC. O modelo conta, no processo

de segmentação, com a delimitação manual de marcos anatômicos ósseos na tíbia, o que limita sua utilização em estudos com muitas imagens. A inclusão de radiografias KL=1 e a verificação da acurácia em uma amostra com maior número de pacientes sem OA (como ocorre na população em geral) seria interessante para verificar a acurácia do modelo em um contexto mais próximo do real. Tiulpin et al. (19) criaram uma rede neural siamesa para a classificação KL automática de radiografias dos joelhos. Foram utilizadas 18376 radiografias do MOST para o treinamento da rede, com 2957 e 5960 imagens da OAI sendo empregadas para a validação e teste. Para estimar o desempenho do modelo para o diagnóstico de OAJ consideraram $KL \geq 2$, tendo alcançado AUC de 0,93. No treinamento, foram utilizados exames seriados dos participantes (de todas as visitas de controle) e todas as angulações de feixe de raio-x disponíveis (5°, 10° e 15°), o que contribuiu para a robustez do modelo.

Algumas técnicas para tratar o desequilíbrio entre as classes foram tentadas, como a realização de *data augmentation* utilizada no presente trabalho apenas sobre os dados positivos e mudança na função de perda (penalizando 10 vezes mais os erros do tipo II), sem repercussões na AUC do modelo. De fato, estudos recentes em dados tabulares (39,40) demonstraram que esses e outros métodos de correção podem até reduzir a AUC, especialmente para modelos com bom desempenho (41).

Analisando os resultados da RNC com base nos diferentes limiares e métricas de desempenho escolhidas, verifica-se a predominância da especificidade sobre a sensibilidade. Esse padrão se deve ao desequilíbrio entre as classes (maior número de exames negativos para OA) e/ou ao número limitado de exames utilizados no treinamento.

A calibração da rede neural por meio da definição de limiares foi o mecanismo que mais impactou nas métricas de desempenho. De fato, a definição de um menor limiar para a definição de OA permitiu incremento na sensibilidade do modelo, enquanto o aumento do limiar ocasionou incremento na especificidade do modelo.

A calibração do modelo deve ser feita de acordo com a aplicação que se pretende dar a ela. Nesse sentido, um modelo com maior acurácia balanceada seria mais apropriado caso sua aplicação seja como uma ferramenta de segunda leitura, enquanto um modelo mais sensível seria preferível para sua utilização como método

de triagem (16). Uma mesma rede neural com dois ou mais limiares ou, ainda, mais de uma rede neural podem ser também utilizadas em conjunto, especialmente quando se considera o baixo custo computacional e de tempo relacionado à utilização de modelos pré-treinados. O diagnóstico de diversas doenças é feito dessa forma: realização de dois testes de forma seriada, um mais sensível e outro mais específico (42). Assim, a avaliação das imagens por um modelo muito sensível, para triar exames possivelmente alterados, poderia ser seguida pela avaliação com um modelo mais balanceado ou mais específico. Outra possibilidade seria a aplicação de dois modelos sensíveis, sendo que a positividade de qualquer um deles já triaria o exame para segunda leitura. Um fator importante a ser considerado é o tempo necessário para cada uma dessas avaliações. A leitura desse volume de radiografias por humanos demanda centenas de horas, enquanto a realizada pelo software se limita a alguns minutos.

Considerando-se a especificidade alcançada pelo modelo, sua aplicação em tarefas como a verificação de possíveis inconsistências (falsos negativos) no banco de dados ou definição de prioridade na fila de exames a serem analisados. A sensibilidade apresentada permite seu uso como possível ferramenta de triagem de exames normais, o que reduziria o volume de exames a serem avaliados pelos radiologistas. Ressalta-se que o diagnóstico de algumas doenças, como é o caso da OAJ, usualmente exige a associação dos achados radiográficos com dados clínico-epidemiológicos, laboratoriais e/ou de outros métodos de imagem, o que não foi testado no presente estudo.

Por se tratar de modelo do tipo *blackbox*, é importante que as conclusões da rede neural sejam pautadas em aspectos considerados relevantes para o diagnóstico, de forma compreensível pelo ser humano (43). Esse fator é chamado de explicabilidade ou interpretabilidade da rede, podendo ser expresso sob a forma de mapas de atenção (ou de saliência), que destacam as regiões da imagem mais relacionadas à predição do modelo (como osteófitos, espaços articulares, esclerose). Ferramentas de explicabilidade da RNC criada ainda estão em desenvolvimento, sendo uma limitação atual do modelo.

O treinamento e validação da rede neural foi realizado com base na interpretação de dois tecnólogos e um radiologista. Isso se deve ao fato de que a leitura de tais

radiografias é demasiado demorada e trabalhosa, sendo impraticável realizar dupla leitura no contexto de um estudo longitudinal com grande amostra como o ELSA-Brasil MSK. A inclusão de exames provenientes de outros estudos longitudinais, com laudos realizados por um comitê de radiologistas, pode contribuir para a robustez da rede, sendo um passo futuro no seu desenvolvimento. Ainda assim, considerando-se a acurácia alcançada, consideramos que o modelo conseguiu aprender a interpretar as radiografias dos joelhos.

As radiografias do ELSA-Brasil MSK foram realizadas por tecnólogos com treinamento, de forma padronizada, em incidência específica e com a utilização de posicionador próprio para a avaliação de OAJ. Na prática médica, entretanto, a incidência historicamente mais utilizada é a anteroposterior com extensão dos joelhos (13), apesar de apresentar menor acurácia para a avaliação de OAJ, não sendo possível extrapolar os resultados para ambientes ambulatoriais ou hospitalares em geral. O desempenho da RNC não foi ainda testado em populações de outros estudos (como OAI ou MOST). Considerando-se que os dados do OAI estão públicos e que os do MOST serão disponibilizados em 2023, a validação externa do modelo será realizada em estudos futuros. Os bancos de dados de tais estudos podem também ser incorporados no processo de treinamento e validação da rede, melhorando sua performance.

Outra limitação está na redução da resolução das imagens no pré-processamento, o que é usual no desenvolvimento de modelos de IA. Apesar de possibilitar maior tempo-efetividade, tal mecanismo pode limitar os resultados pela perda de informações sutis dos exames.

Em trabalhos futuros, será avaliada a possibilidade da utilização de duas redes neurais simultaneamente, triando para a leitura humana a positividade em qualquer um dos testes, como uma forma de contornar a sensibilidade limitada e permitir sua aplicação ao contexto do ELSA-Brasil MSK. Melhorias na rede neural apresentada, a partir da utilização de imagens das ondas subseqüentes do ELSA-Brasil MSK, bem como de bancos de imagens de outros estudos (como do OAI e MOST), devem contribuir para o aumento do desempenho e robustez da rede. Tais aprimoramentos estão sendo implementados e serão publicados em trabalhos futuros.

Conclusão

A RNC desenvolvida apresenta desempenho compatível com a das redes neurais treinadas com radiografias de estudos internacionais. A acurácia e AUC alcançadas permitem sua utilização como ferramenta de segunda leitura no ELSA-Brasil MSK, contribuindo para contornar a limitada disponibilidade de radiologistas treinados na interpretação de radiografias do joelho.

A validação do modelo em populações diferentes da do ELSA-Brasil MSK, tanto de outros estudos longitudinais quanto da prática clínica, é importante para a sua adoção futura. O código-fonte será disponibilizado publicamente, de forma a permitir que essa validação externa ocorra em estudos futuros, desta ou de outras equipes.

Agradecimentos

O ELSA-Brasil é financiado pelo Ministério da Saúde (Departamento de Ciência e Tecnologia) e pelo Ministério de Ciência e Tecnologia (Financiadora de Estudos e Projetos e CNPq-BR) do Brasil, 01 06 0010.00 RS, 01 06 0212.00 BA, 01 06 0300.00 ES, 01 06 0278.00 MG, 01 06 0115.00 SP, 01 06 0071.00 RJ. O ELSA-Brasil MSK recebe apoio e agradece o recebimento de fomento dos seguintes órgãos/fundações: CAPES-BR (post-doc/SUS 054/2010), FAPEMIG-BR (APQ-00921-16, APQ-00549-22) e CNPq-BR (423585/2016-9, 404728/2021-9).

Referências

1. Santo L, Okeyode T. National Ambulatory Medical Care Survey: 2018 National Summary Tables [Internet]. 2018 [citado 2022 mar 3]. Available from: https://www.cdc.gov/nchs/data/ahcd/namcs_summary/2018-namcs-web-tables-508.pdffrom: https://www.cdc.gov/nchs/data/ahcd/namcs_summary/2018-namcs-web-tables-508.pdf
2. Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PWF, et al. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. *Am J Public Health*. 1994;84(3):351.
3. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*. 2012;380:2163–96.
4. Neogi T, Zhang Y. Epidemiology of osteoarthritis. *Rheum Dis Clin North Am*. 2013 fev;39(1):1–19.
5. Oliveria SA, Felson DT, Reed JI, Cirillo PA, Walker AM. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. *Arthritis Rheum*. 1995 ago 1;38(8):1134–41.
6. Telles RW, Machado LAC, Costa-Silva L, Barreto SM. Cohort Profile Update: The Brazilian Longitudinal Study of Adult Health Musculoskeletal (ELSA-Brasil MSK) cohort. *Int J Epidemiol*. 2022 abr 26.
7. Sharma L. Osteoarthritis of the Knee. Solomon CG, organizador. *N Engl J Med*. 2021 jan 6;384(1):51–9.
8. Hunter DJ, Schofield D, Callander E. The individual and socioeconomic impact of osteoarthritis. *Nature Reviews Rheumatology* 2014 10:7. 2014 mar 25;10(7):437–41.
9. March LM, Bachmeier CJM. 10 Economics of osteoarthritis: a global perspective. *Baillieres Clin Rheumatol*. 1997 nov 1;11(4):817–34.

10. Hunter DJ, Schofield D, Callander E. The individual and socioeconomic impact of osteoarthritis. *Nat Rev Rheumatol*. 2014;10(7):437–41.
11. Pereira D, Peleteiro B, Araújo J, Branco J, Santos RA, Ramos E. The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. *Osteoarthritis Cartilage*. 2011 nov 1;19(11):1270–85.
12. Miguel R de CC, Machado LA, Costa-Silva L, Telles RW, Barreto SM. Performance of distinct knee osteoarthritis classification criteria in the ELSA-Brasil musculoskeletal study. *Clin Rheumatol*. 2019 mar 14;38(3):793–802.
13. Buckland-Wright C. Which radiographic techniques should we use for research and clinical practice? *Best Pract Res Clin Rheumatol*. 2006 fev 1;20(1):39–55.
14. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. 1957;16(4):494–502.
15. Binvignat M, Pedoia V, Butte AJ, Louati K, Klatzmann D, Berenbaum F, et al. Use of machine learning in osteoarthritis research: a systematic literature review. *RMD Open*. 2022 mar 1;8(1):e001998.
16. Chartrand G, Cheng PM, Vorontsov E, Drozdal M, Turcotte S, Pal CJ, et al. Deep learning: A primer for radiologists. *Radiographics*. 2017 nov 1;37(7):2113–31.
17. Yeoh PSQ, Lai KW, Goh SL, Hasikin K, Hum YC, Tee YK, et al. Emergence of Deep Learning in Knee Osteoarthritis Diagnosis. *Comput Intell Neurosci*. 2021.
18. Thomas KA, Kidziński Ł, Halilaj E, Fleming SL, Venkataraman GR, Oei EHG, et al. Automated Classification of Radiographic Knee Osteoarthritis Severity Using Deep Neural Networks. *Radiol Artif Intell*. 2020 mar 1;2(2):e190065.
19. Tiulpin A, Thevenot J, Rahtu E, Lehenkari P, Saarakkala S. Automatic Knee Osteoarthritis Diagnosis from Plain Radiographs: A Deep Learning-Based Approach. *Sci Rep*. 2018 dez 1;8(1).
20. Abedin J, Antony J, McGuinness K, Moran K, O'Connor NE, Rebholz-Schuhmann D, et al. Predicting knee osteoarthritis severity: comparative modeling

based on patient's data and plain X-ray images. *Scientific Reports* 2019 9:1 [Internet]. 2019 abr 8;9(1):1–11.

21. Norman B, Pedoia V, Noworolski A, Link TM, Majumdar S. Applying Densely Connected Convolutional Neural Networks for Staging Osteoarthritis Severity from Plain Radiographs. *J Digit Imaging*. 2019 jun 15;32(3):471.
22. Chen P, Gao L, Shi X, Allen K, Yang L. Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss. *Computerized Medical Imaging and Graphics*. 2019 jul 1;75:84–92.
23. Swiecicki A, Li N, O'Donnell J, Said N, Yang J, Mather RC, et al. Deep learning-based algorithm for assessment of knee osteoarthritis severity in radiographs matches performance of radiologists. *Comput Biol Med*. 2021 jun 1;133:104334.
24. Brahim A, Jennane R, Riad R, Janvier T, Khedher L, Toumi H, et al. A decision support tool for early detection of knee OsteoArthritis using X-ray imaging and machine learning: Data from the OsteoArthritis Initiative. *Computerized Medical Imaging and Graphics*. 2019 abr 1;73:11–8.
25. Schmidt MI, Duncan BB, Mill JG, Lotufo PA, Chor D, Barreto SM, et al. Cohort Profile: Longitudinal Study of Adult Health (ELSA-Brasil). *Int J Epidemiol*. 2015 fev 1;44(1):68–75.
26. Machado LAC, Costa-Silva L, Barreto SM, Telles RW, Naves EB, Carneiro AMFC. Positioner for the acquisition and quality control of radiographic images of knees in the fixed flexion position (radiolfixer). Brazil: National Institute of Industrial Property; BR102013033625-4, 2013. p. 161–161.
27. Telles RW, Costa-Silva L, Machado LAC, Reis RCP dos, Barreto SM. Fixed-flexion knee radiography using a new positioning device produced highly repeatable measurements of joint space width: ELSA-Brasil Musculoskeletal Study (ELSA-Brasil MSK). *Rev Bras Reumatol*. 2017;57(2):154–61.
28. Nevitt MC, Felson DT, Lester G. The Osteoarthritis Initiative. Protocol for the cohort study. 2017 out.

29. Segal NA, Nevitt MC, Gross KD, Hietpas J, Glass NA, Lewis CE, et al. The Multicenter Osteoarthritis Study (MOST): Opportunities for Rehabilitation Research. *PM&R [Internet]*. 2013 ago;5(8):647–54.
30. Deng J, Dong W, Socher R, Li LJ, Kai Li, Li Fei-Fei. ImageNet: A large-scale hierarchical image database. 2010 mar 1;248–55.
31. DenseNet | Papers With Code [Internet]. [citado 2022 set 20]. Available from: <https://paperswithcode.com/model/densenet?variant=densenet-161>
32. Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. 2017 nov 6;2017-January:2261–9.
33. Gupta A, Tatbul N, Marcus R, Zhou S, Lee I, Gottschlich J. Class-Weighted Evaluation Metrics for Imbalanced Data Classification. 2020 out 12.
34. Araújo DC, Veloso AA, Borges KBG, Carvalho M das G. Prognosing the risk of COVID-19 death through a machine learning-based routine blood panel: A retrospective study in Brazil. *Int J Med Inform*. 2022 set 1;165:104835.
35. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett*. 2006 jun 1;27(8):861–74.
36. Carter J v., Pan J, Rai SN, Galandiuk S. ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. *Surgery*. 2016 jun 1;159(6):1638–45.
37. Ling CX, Huang J, Zhang H. AUC: A better measure than accuracy in comparing learning algorithms. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2003;2671:329–41.
38. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit*. 1997 jul 1;30(7):1145–59.
39. van den Goorbergh R, van Smeden M, Timmerman D, van Calster B. The harm of class imbalance corrections for risk prediction models: illustration and

simulation using logistic regression. *Journal of the American Medical Informatics Association*. 2022 jun 10.

40. Elor Y, Averbuch-Elor H. To SMOTE, or not to SMOTE? 2022 jan 21.

41. Araújo DC, Veloso AA, Borges KBG, Carvalho M das G. Prognosing the risk of COVID-19 death through a machine learning-based routine blood panel: A retrospective study in Brazil. *Int J Med Inform*. 2022 set 1;165:104835.

42. Weinstein S, Obuchowski NA, Lieber ML. Clinical evaluation of diagnostic tests. *AJR Am J Roentgenol*. 2005;184(1):14–9.

43. Fan FL, Xiong J, Li M, Wang G. On Interpretability of Artificial Neural Networks: A Survey. *IEEE Trans Radiat Plasma Med Sci*. 2021 nov 1;5(6):741–60.

5 CONSIDERAÇÕES FINAIS

Algoritmos de inteligência artificial apresentam-se como promissoras ferramentas de auxílio ao diagnóstico para diversas doenças. O presente estudo verificou o desempenho de diferentes RNC para o diagnóstico radiográfico de OAJ.

No presente estudo, verificou-se o desempenho de uma rede neural desenvolvida com base em radiografias do OAI para o diagnóstico de OAJ nos exames do ELSA-Brasil MSK. A performance do modelo é considerada excelente (CARTER et al., 2016), com AUC de 0,901 (IC 95% 0,858-0,945). A aplicação do modelo de Thomas et al. (2020) a uma população com perfil epidemiológico diferente da que foi treinado, sem ajustes nos parâmetros ou retreino, aponta para a robustez da RNC, sendo especialmente importante quando consideramos a inerente natureza *blackbox* de redes neurais.

O presente estudo desenvolveu também uma RNC com base nas radiografias da linha de base do ELSA-Brasil MSK. A RNC proposta apresentou AUC 0,866 (IC 95% 0,842-0,882), sem diferença estatística para o desempenho do modelo de Thomas et al. (2020).

Ambas as redes neurais avaliadas podem ser utilizadas como ferramentas de triagem, de modo a reduzir o volume total de exames a serem lidos pelo radiologista, mas não de forma isolada. Em trabalhos futuros, será avaliada a possibilidade da utilização de mais de uma rede neural simultaneamente, triando para a leitura humana a positividade em qualquer um dos testes. Ambas as RNC podem, considerando-se a acurácia e especificidade alcançadas, ser utilizadas para verificação de possíveis inconsistências no banco de dados do ELSA-Brasil MSK, definição de prioridade na fila de exames a serem analisados e seleção de exames para segunda leitura.

O desempenho das redes neurais é insuficiente para que a interpretação humana especializada seja dispensada. Cabe ressaltar ainda que o diagnóstico de algumas doenças, como é o caso da OAJ, usualmente exige a associação dos achados radiográficos com dados clínico-epidemiológicos, laboratoriais e/ou de outros métodos de imagem, o que não foi testado no presente estudo.

Em trabalhos futuros pretende-se verificar a consistência e coerência de redes neurais em estudo seriados, o que é de grande relevância para estudos longitudinais como o

ELSA-Brasil MSK. A utilização de dados clínico-epidemiológicos associados às radiografias, visando à estimativa de risco de se desenvolver OAJ ao longo da vida, será também alvo de pesquisas futuras.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- ABBAFATI, C. et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. **Lancet (London, England)**, v. 396, n. 10258, p. 1204, 17 out. 2020.
- ALLEN, K. D.; THOMA, L. M.; GOLIGHTLY, Y. M. Epidemiology of osteoarthritis. **Osteoarthritis and Cartilage**, v. 30, n. 2, p. 184–195, 1 fev. 2022.
- ALTMAN, R. et al. Development of criteria for the classification and reporting of osteoarthritis: Classification of osteoarthritis of the knee. **Arthritis & Rheumatism**, v. 29, n. 8, p. 1039–1049, 1 ago. 1986.
- ALTMAN, R. D. Criteria for classification of clinical osteoarthritis. **The Journal of rheumatology. Supplement**, v. 27, n. SUPPL. 27, p. 10–12, 1 fev. 1991.
- ALTMAN, R. D.; GOLD, G. E. Atlas of individual radiographic features in osteoarthritis, revised. **Osteoarthritis and Cartilage**, v. 15, p. A1–A56, 2007.
- AMIN, S. et al. The relationship between cartilage loss on magnetic resonance imaging and radiographic progression in men and women with knee osteoarthritis. **Arthritis and rheumatism**, v. 52, n. 10, p. 3152–3159, out. 2005.
- AQUINO, E. M. L. et al. Brazilian Longitudinal Study of Adult Health (ELSA-Brasil): objectives and design. **American journal of epidemiology**, v. 175, n. 4, p. 315–324, 15 fev. 2012.
- BACKHAUS, M. et al. Guidelines for musculoskeletal ultrasound in rheumatology. **Annals of the Rheumatic Diseases**, v. 60, n. 7, p. 641–649, 1 jul. 2001.
- BARBOUR, K. E. et al. Meeting physical activity guidelines and the risk of incident knee osteoarthritis: a population-based prospective cohort study. **Arthritis care & research**, v. 66, n. 1, p. 139–146, jan. 2014.
- BIJLSMA, J. W. J.; BERENBAUM, F.; LAFEBER, F. P. J. G. Osteoarthritis: an update with relevance for clinical practice. **Lancet (London, England)**, v. 377, n. 9783, p. 2115–2126, 2011.
- BINVIGNAT, M. et al. Use of machine learning in osteoarthritis research: a systematic literature review. **RMD Open**, v. 8, n. 1, p. e001998, 1 mar. 2022.
- BLANCO, F. J. et al. Prevalence of symptomatic osteoarthritis in Spain: EPISER2016 study. **Reumatología Clínica (English Edition)**, v. 17, n. 8, p. 461–470, 1 out. 2021.
- BLUEMKE, D. A. et al. Assessing radiology research on artificial intelligence: A brief guide for authors, reviewers, and readers-from the Radiology Editorial Board. **Radiology**, v. 294, n. 2, p. 487–489, 31 dez. 2020.
- BREJNEBØL, M. W. et al. External validation of an artificial intelligence tool for radiographic knee osteoarthritis severity classification. **European Journal of Radiology**, v. 150, p. 110249, 1 maio 2022.

BUCKLAND-WRIGHT, C. Which radiographic techniques should we use for research and clinical practice? **Best Practice & Research Clinical Rheumatology**, v. 20, n. 1, p. 39–55, 1 fev. 2006.

BUCKLAND-WRIGHT, J. C. Quantitative radiography of osteoarthritis. **Annals of the Rheumatic Diseases**, v. 53, n. 4, p. 268–275, 1 abr. 1994.

BUCKLAND-WRIGHT, J. C. et al. Accuracy and precision of joint space width measurements in standard and macroradiographs of osteoarthritic knees. **Annals of the rheumatic diseases**, v. 54, n. 11, p. 872–880, 1995.

CARTER, J. V. et al. ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. **Surgery**, v. 159, n. 6, p. 1638–1645, 1 jun. 2016.

CHAISSON, C. E. et al. Detecting radiographic knee osteoarthritis: what combination of views is optimal? **Rheumatology (Oxford, England)**, v. 39, n. 11, p. 1218–1221, 2000.

CHARTRAND, G. et al. Deep learning: A primer for radiologists. **Radiographics**, v. 37, n. 7, p. 2113–2131, 1 nov. 2017.

CISTERNAS, M. G. et al. Alternative Methods for Defining Osteoarthritis and the Impact on Estimating Prevalence in a US Population-Based Survey. **Arthritis care & research**, v. 68, n. 5, p. 574, 1 maio 2016.

DOI, K. Computer-Aided Diagnosis in Medical Imaging: Historical Review, Current Status and Future Potential. **Comput Med Imaging Graph**, v. 31, n. 4–5, p. 198, 2007.

DRIBAN, J. B. et al. Is Participation in Certain Sports Associated With Knee Osteoarthritis? A Systematic Review. **Journal of Athletic Training**, v. 52, n. 6, p. 497–506, 2 jun. 2017.

EMRANI, P. S. et al. Joint space narrowing and Kellgren–Lawrence progression in knee osteoarthritis: an analytic literature synthesis. **Osteoarthritis and Cartilage**, v. 16, n. 8, p. 873–882, 1 ago. 2008.

FELSON, D. T. et al. The prevalence of knee osteoarthritis in the elderly. the framingham osteoarthritis study. **Arthritis & Rheumatism**, v. 30, n. 8, p. 914–918, 1 ago. 1987.

FELSON, D. T. The epidemiology of knee osteoarthritis: Results from the framingham osteoarthritis study. **Seminars in Arthritis and Rheumatism**, v. 20, n. 3, p. 42–50, 1 dez. 1990.

FELSON, D. T. et al. The incidence and natural history of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. **Arthritis and rheumatism**, v. 38, n. 10, p. 1500–1505, 1995.

FELSON, D. T. et al. Effect of recreational physical activities on the development of knee osteoarthritis in older adults of different weights: the Framingham Study. **Arthritis and rheumatism**, v. 57, n. 1, p. 6–12, 15 fev. 2007.

FELSON, D. T. et al. A New Approach Yields High Rates of Radiographic Progression in Knee Osteoarthritis. **The Journal of rheumatology**, v. 35, n. 10, p. 2047, out. 2008.

FELSON, D. T. et al. Defining radiographic incidence and progression of knee osteoarthritis: suggested modifications of the Kellgren and Lawrence scale. **Annals of the rheumatic diseases**, v. 70, n. 11, p. 1884–1886, nov. 2011.

FREDERICK, W.; LANE E, N.; BUCKLAND-WRIGHT, C. Radiographic Methods in Knee Osteoarthritis: A Further Comparison of Semiflexed (MTP), Schuss-Tunnel, and Weight-Bearing Anteroposterior Views for Joint Space Narrowing and Osteophytes. **J Rheumatol.**, v. 12, p. 2597–2601, 2002.

FREIBERG, A.; BOLM-AUDORFF, U.; SEIDLER, A. The Risk of Knee Osteoarthritis in Professional Soccer Players: A Systematic Review with Meta-Analyses. **Deutsches Ärzteblatt International**, v. 118, n. 4, p. 49, 30 out. 2021.

GUERMAZI, A. et al. Prevalence of abnormalities in knees detected by MRI in adults without knee osteoarthritis: population based observational study (Framingham Osteoarthritis Study). **BMJ**, v. 345, n. 7874, 29 ago. 2012.

HUNTER, D. J. et al. Change in joint space width: hyaline articular cartilage loss or alteration in meniscus? **Arthritis and rheumatism**, v. 54, n. 8, p. 2488–2495, ago. 2006.

HUNTER, D. J. et al. Imaging biomarker validation and qualification report: Sixth oarsi workshop on imaging in osteoarthritis combined with third oa biomarkers workshop. **Osteoarthritis and Cartilage**, v. 21, n. 7, p. 939–942, 1 jul. 2013.

HUNTER, D. J.; BIERMA-ZEINSTRA, S. Osteoarthritis. **The Lancet**, v. 393, n. 10182, p. 1745–1759, 27 abr. 2019.

HUNTER, D. J.; SCHOFIELD, D.; CALLANDER, E. The individual and socioeconomic impact of osteoarthritis. **Nature reviews. Rheumatology**, v. 10, n. 7, p. 437–441, 2014.

JIN, Z. et al. Global, regional and national trends in sex- and age-specific disability-adjusted life years of musculoskeletal disorders, 1990–2019. **Rheumatology**, 27 out. 2021.

JONES, G. et al. Early radiographic osteoarthritis is associated with substantial changes in cartilage volume and tibial bone surface area in both males and females. **Osteoarthritis and Cartilage**, v. 12, n. 2, p. 169–174, 2004.

KATZ, J. N.; ARANT, K. R.; LOESER, R. F. Diagnosis and Treatment of Hip and Knee Osteoarthritis: A Review. **JAMA**, v. 325, n. 6, p. 568–578, 9 fev. 2021.

KEEN, H. I.; WAKEFIELD, R. J.; CONAGHAN, P. G. A systematic review of ultrasonography in osteoarthritis. **Annals of the Rheumatic Diseases**, v. 68, n. 5, p. 611–619, 1 maio 2009.

KELLGREN, J. H.; LAWRENCE, J. S. Radiological assessment of osteo-arthrosis. **Annals of the rheumatic diseases**, v. 16, n. 4, p. 494–502, 1957.

KIM, D. H. et al. Can Additional Patient Information Improve the Diagnostic Performance of Deep Learning for the Interpretation of Knee Osteoarthritis Severity. **Journal of Clinical Medicine** 2020, Vol. 9, Page 3341, v. 9, n. 10, p. 3341, 18 out. 2020.

KOENIGKAM SANTOS, M. et al. Artificial intelligence, machine learning, computer-aided diagnosis, and radiomics: Advances in imaging towards to precision medicine. **Radiologia Brasileira**, v. 52, n. 6, p. 387–396, 1 nov. 2019.

KOTHARI, M. et al. Fixed-flexion radiography of the knee provides reproducible joint space width measurements in osteoarthritis. **European Radiology**, v. 14, n. 9, p. 1568–1573, 19 set. 2004.

KRAUS, V. B. et al. Call for Standardized Definitions of Osteoarthritis and Risk Stratification for Clinical Trials and Clinical Use. **Osteoarthritis and cartilage**, v. 23, n. 8, p. 1233, 1 ago. 2015.

LEE, L. S. et al. Imaging of knee osteoarthritis: A review of current evidence and clinical guidelines. **Musculoskeletal Care**, v. 19, n. 3, p. 363–374, 1 set. 2021.

LESTER, G. The Osteoarthritis Initiative: A NIH Public–Private Partnership. **HSS Journal**, v. 8, n. 1, p. 62, 23 fev. 2012.

LEYLAND, K. M. et al. Harmonising measures of knee and hip osteoarthritis in population-based cohort studies: an international study. **Osteoarthritis and cartilage**, v. 26, n. 7, p. 872–879, 1 jul. 2018.

LI, D. et al. The Prevalence of Symptomatic Knee Osteoarthritis in Relation to Age, Sex, Area, Region, and Body Mass Index in China: A Systematic Review and Meta-Analysis. **Frontiers in Medicine**, v. 7, p. 304, 16 jul. 2020.

LIU, B.; LUO, J.; HUANG, H. Toward automatic quantification of knee osteoarthritis severity using improved Faster R-CNN. **Int J Comput Assist Radiol Surg**, v. 15, n. 3, p. 457–466, 1 mar. 2020.

LO VERCIO, L. et al. Supervised machine learning tools: a tutorial for clinicians. **Journal of Neural Engineering**, v. 17, n. 6, p. 062001, 19 nov. 2020.

LOESER, R. F. et al. Osteoarthritis: A disease of the joint as an organ. **Arthritis & Rheumatism**, v. 64, n. 6, p. 1697–1707, 1 jun. 2012.

LOESER, R. F.; COLLINS, J. A.; DIEKMAN, B. O. Ageing and the pathogenesis of osteoarthritis. **Nature Reviews Rheumatology** 2016 12:7, v. 12, n. 7, p. 412–420, 19 maio 2016.

LOSINA, E. et al. Lifetime risk and age at diagnosis of symptomatic knee osteoarthritis in the US. **Arthritis Care and Research**, v. 65, n. 5, p. 703–711, 1 maio 2013.

MACHADO, L. A. C. et al. **Positioner for the acquisition and quality control of radiographic images of knees in the fixed flexion position (radiolixer).**

Brazil National Institute of Industrial Property, , 27 dez. 2013.

MARCH, L. M.; BACHMEIER, C. J. M. 10 Economics of osteoarthritis: a global perspective. **Baillière's Clinical Rheumatology**, v. 11, n. 4, p. 817–834, 1 nov. 1997.

MATHIESSEN, A. et al. Imaging of osteoarthritis (OA): What is new? **Best Practice & Research Clinical Rheumatology**, v. 30, n. 4, p. 653–669, 1 ago. 2016.

MAZZUCA, S. A.; BRANDT, K. D.; KATZ, B. P. Is conventional radiography suitable for evaluation of a disease-modifying drug in patients with knee osteoarthritis? **Osteoarthritis and Cartilage**, v. 5, n. 4, p. 217–226, 1 jul. 1997.

MENASHE, L. et al. The diagnostic performance of MRI in osteoarthritis: a systematic review and meta-analysis. **Osteoarthritis and Cartilage**, v. 20, n. 1, p. 13–21, 1 jan. 2012.

MESSIEH, S. S.; FOWLER, P. J.; MUNRO, T. Anteroposterior radiographs of the osteoarthritic knee. **The Journal of bone and joint surgery. British volume**, v. 72, n. 4, p. 639–640, 1990.

MIGUEL, R. DE C. C. et al. Performance of distinct knee osteoarthritis classification criteria in the ELSA-Brasil musculoskeletal study. **Clinical Rheumatology**, v. 38, n. 3, p. 793–802, 14 mar. 2019.

MURPHY, L. et al. Lifetime risk of symptomatic knee osteoarthritis. **Arthritis Care & Research**, v. 59, n. 9, p. 1207–1213, 15 set. 2008.

MURRAY, C. J. L. et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990-2013: quantifying the epidemiological transition. **Lancet (London, England)**, v. 386, n. 10009, p. 2145–2191, 28 nov. 2015.

NEOGI, T. et al. Association between radiographic features of knee osteoarthritis and pain: results from two cohort studies. **BMJ**, v. 339, n. 7719, p. 498–501, 21 ago. 2009.

NEOGI, T.; ZHANG, Y. Epidemiology of osteoarthritis. **Rheumatic diseases clinics of North America**, v. 39, n. 1, p. 1–19, fev. 2013.

NEVALAINEN, M. T. et al. Ultrasonography of the late-stage knee osteoarthritis prior to total knee arthroplasty: comparison of the ultrasonographic, radiographic and intra-operative findings. **Scientific Reports**, v. 8, n. 1, p. 1–7, 10 dez. 2018.

NEVITT, M. C.; FELSON, D. T.; LESTER, G. **The Osteoarthritis Initiative. Protocol for the cohort study.** [s.l: s.n.]. Disponível em: <<https://nda.nih.gov/static/docs/StudyDesignProtocolAndAppendices.pdf>>. Acesso em: 24 mar. 2022.

NIU, J. et al. Metabolic Syndrome, Its Components, and Knee Osteoarthritis: The Framingham Osteoarthritis Study. **Arthritis & Rheumatology**, v. 69, n. 6, p. 1194–1203, 1 jun. 2017.

ODDING, E. et al. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. **Annals of the rheumatic diseases**, v. 57, n. 4, p. 203–208, 1998.

OLIVERIA, S. A. et al. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. **Arthritis & Rheumatism**, v. 38, n. 8, p. 1134–1141, 1 ago. 1995.

OO, W. M.; BO, M. T. Role of ultrasonography in knee osteoarthritis. **Journal of Clinical Rheumatology**, v. 22, n. 6, p. 324–329, 1 set. 2016.

PEAT, G. et al. Relative prevalence and distribution of knee, hand and foot symptomatic osteoarthritis subtypes in an English population. **Musculoskeletal Care**, v. 18, n. 2, p. 219–224, 1 jun. 2020.

PEREIRA, D. et al. The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. **Osteoarthritis and Cartilage**, v. 19, n. 11, p. 1270–1285, 1 nov. 2011.

PETERFY, C. et al. Comparison of fixed-flexion positioning with fluoroscopic semi-flexed positioning for quantifying radiographic joint-space width in the knee: test-retest reproducibility. **Skeletal Radiology**, v. 32, n. 3, p. 128–132, 1 mar. 2003.

PODLIPSKÁ, J. et al. Comparison of Diagnostic Performance of Semi-Quantitative Knee Ultrasound and Knee Radiography with MRI: Oulu Knee Osteoarthritis Study. **Scientific Reports**, v. 6, n. 1, p. 1–12, 1 mar. 2016.

RAVAUD, P. et al. Knee joint space width measurement: an experimental study of the influence of radiographic procedure and joint positioning. **British journal of rheumatology**, v. 35, n. 8, p. 761–766, 1996.

RESNICK, D.; VINT, V. The “Tunnel” view in assessment of cartilage loss in osteoarthritis of the knee. **Radiology**, v. 137, n. 2, p. 547–548, 1 nov. 1980.

SAFIRI, S. et al. Global, regional and national burden of osteoarthritis 1990-2017: a systematic analysis of the Global Burden of Disease Study 2017. **Annals of the Rheumatic Diseases**, v. 79, n. 6, p. 819–828, 1 jun. 2020.

SANTO, L.; OKEYODE, T. **National Ambulatory Medical Care Survey: 2018 National Summary Tables**. Disponível em: <https://www.cdc.gov/nchs/data/ahcd/namcs_summary/2018-namcs-web-tables-508.pdf>from: https://www.cdc.gov/nchs/data/ahcd/namcs_summary/2018-namcs-web-tables-508.pdf>. Acesso em: 3 mar. 2022.

SCHMIDT, M. I. et al. Cohort Profile: Longitudinal Study of Adult Health (ELSA-Brasil). **International journal of epidemiology**, v. 44, n. 1, p. 68–75, 1 fev. 2015.

- SCHOFIELD, D. J. et al. The personal and national costs of lost labour force participation due to arthritis: An economic study. **BMC Public Health**, v. 13, n. 1, p. 1–10, 3 mar. 2013.
- SEGAL, N. A. et al. The Multicenter Osteoarthritis Study (MOST): Opportunities for Rehabilitation Research. **PM&R**, v. 5, n. 8, p. 647–654, ago. 2013.
- SHARMA, L. Osteoarthritis of the Knee. **N Engl J Med**, v. 384, n. 1, p. 51–59, 6 jan. 2021.
- SILVERWOOD, V. et al. Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. **Osteoarthritis and Cartilage**, v. 23, n. 4, p. 507–515, 1 abr. 2015.
- SOHN, D. H. et al. Plasma proteins present in osteoarthritic synovial fluid can stimulate cytokine production via Toll-like receptor 4. **Arthritis Research and Therapy**, v. 14, n. 1, p. 1–13, 8 jan. 2012.
- SOWERS, M. F. et al. Knee osteoarthritis in obese women with cardiometabolic clustering. **Arthritis and rheumatism**, v. 61, n. 10, p. 1328–1336, 2009.
- TALJANOVIC, M. S. et al. Bone marrow edema pattern in advanced hip osteoarthritis: Quantitative assessment with magnetic resonance imaging and correlation with clinical examination, radiographic findings, and histopathology. **Skeletal Radiology**, v. 37, n. 5, p. 423–431, 1 maio 2008.
- TELLES, R. W. et al. Fixed-flexion knee radiography using a new positioning device produced highly repeatable measurements of joint space width: ELSA-Brasil Musculoskeletal Study (ELSA-Brasil MSK). **Revista Brasileira de Reumatologia**, v. 57, n. 2, p. 154–161, 2017.
- TELLES, R. W. et al. Cohort Profile Update: The Brazilian Longitudinal Study of Adult Health Musculoskeletal (ELSA-Brasil MSK) cohort. **International journal of epidemiology**, 26 abr. 2022.
- THOMAS, K. A. et al. Automated Classification of Radiographic Knee Osteoarthritis Severity Using Deep Neural Networks. **Radiology. Artificial intelligence**, v. 2, n. 2, p. e190065, 1 mar. 2020.
- VAN DEN BOSCH, M. H. J.; VAN LENT, P. L. E. M.; VAN DER KRAAN, P. M. Identifying effector molecules, cells, and cytokines of innate immunity in OA. **Osteoarthritis and Cartilage**, v. 28, n. 5, p. 532–543, 1 maio 2020.
- VOS, T. et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. **The Lancet**, v. 380, p. 2163–2196, 2012.
- WALLACE, I. J. et al. Knee osteoarthritis has doubled in prevalence since the mid-20th century. **Proc Natl Acad Sci USA**, v. 114, n. 35, p. 9332–9336, 29 ago. 2017.

WANG, X.; OO, W. M.; LINKLATER, J. M. What is the role of imaging in the clinical diagnosis of osteoarthritis and disease management? **Rheumatology**, v. 57, n. suppl_4, p. iv51–iv60, 1 maio 2018.

WOOD, M. J. et al. Macrophage proliferation distinguishes 2 subgroups of knee osteoarthritis patients. **JCI Insight**, v. 4, n. 2, 1 fev. 2019.

YEOH, P. S. Q. et al. Emergence of Deep Learning in Knee Osteoarthritis Diagnosis. **Computational Intelligence and Neuroscience**, v. 2021, 2021.

ZARRINGAM, D.; SARIS, D. B. F.; BEKKERS, J. E. J. The Value of SPECT/CT for Knee Osteoarthritis: A Systematic Review. **Cartilage**, v. 12, n. 4, p. 431–437, 1 out. 2021.

ZHANG, W. et al. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. **Annals of the Rheumatic Diseases**, v. 69, n. 3, p. 483–489, 1 mar. 2010.

ZHANG, Y.; JORDAN, J. M. Epidemiology of Osteoarthritis. **Clinics in Geriatric Medicine**, v. 26, n. 3, p. 355–369, 1 ago. 2010.

7 APÊNDICES

I. Acesso ao código do modelo desenvolvido (Kneelsa)

A rede neural convolucional aqui apresentada está em desenvolvimento e será disponibilizada após a sua conclusão. O estado atual pode ser acessado em:

<https://github.com/jgdjulio/kneelsa>

8 ANEXOS

I. Trabalho apresentado no 51º Congresso Brasileiro de Radiologia e Diagnóstico por Imagem

Resumo

Página Inicial » Inscrições Científicas » Trabalhos

Dados do Trabalho

Título

Desenvolvimento de rede neural convolucional para o diagnóstico radiográfico de osteoartrite dos joelhos no Estudo Longitudinal da Saúde do Adulto – Musculosquelético (ELSA-Brasil MSK)

Introdução

A osteoartrite dos joelhos (OAJ) é uma das doenças mais prevalentes na população mundial, associada a elevada morbidade. Seu diagnóstico é frequentemente realizado a partir da avaliação de radiografias. Em estudos longitudinais de larga escala, esse processo é laborioso e demorado, exigindo treinamento específico. A aplicação de método de classificação automática pode contribuir para o fluxo de diagnóstico, reduzindo a quantidade de exames a serem avaliados.

A rede neural convolucional (RNC) é um dos modelos mais utilizados para a classificação de imagens médicas. Seu desenvolvimento está concentrado em estudos nos EUA e na Europa, sendo escassas as ferramentas validadas em outras populações. Revisões recentes não identificaram abrangência da população brasileira na validação das atuais RNCs para diagnóstico de OAJ.

Este estudo visa ao desenvolvimento de RNC utilizando exames do ELSA-Brasil MSK. Considerando a revisão realizada, trata-se da primeira RNC criada com esse objetivo no Brasil.

Casuística e Métodos

No desenvolvimento da RNC foram utilizados apenas os exames da linha de base do ELSA-Brasil MSK (radiografias bilaterais dos joelhos de 2830 participantes em incidência posteroanterior). Os exames foram interpretados por médico radiologista com treinamento específico.

A rede neural proposta utiliza arquitetura convolucional densamente conectada pré-treinada. As imagens passaram por data augmentation, após a segmentação das regiões de interesse. Foi realizada validação cruzada em dois folds, testando até 12 épocas.

Resultados

A RNC apresentou acurácia de 0,814, com sensibilidade de 0,755 e especificidade de 0,821, com área sob a curva de 0.866 (IC 95% 0.842-0.882). O modelo pode ser otimizado para alcançar, não simultaneamente, valores máximos de 0,907 para acurácia, 0,938 para sensibilidade e 0,994 para especificidade.

Conclusões

A RNC apresentou acurácia próxima a de modelos internacionais. O modelo apresentou alta especificidade, podendo ser utilizado para tarefas determinadas, como a verificação de inconsistências no banco de dados. A sensibilidade, todavia, limita seu uso como ferramenta de triagem.

O desenvolvimento de RNC própria permite a definição de limiares de acordo com a proposta de utilização, priorizando sensibilidade ou especificidade, por exemplo. Um maior volume de imagens de pacientes com OAJ e modificações nos parâmetros do modelo podem contribuir para seu aperfeiçoamento, possibilitando sua aplicação no contexto do ELSA-Brasil MSK e de outros estudos de larga escala.

Palavras Chave

Osteoartrite; joelhos; redes neurais; inteligência artificial; radiografia.

Área

Inteligência Artificial, Inovação e Telerradiologia

Instituições

Universidade Federal de Minas Gerais - Minas Gerais - Brasil

Autores

Júlio Guerra Domingues, Daniella Castro Araújo, Luciana Costa-Silva, Alexei Manso Corrêa Machado, Luciana Andrade Carneiro Machado, Adriano Alonso Veloso, Sandhi Maria Barreto, Rosa Weiss Telles

Certificado de apresentação



CBR22
51º CONGRESSO BRASILEIRO
DE RADIOLOGIA E
DIAGNÓSTICO POR IMAGEM

 **cbr** | Colégio Brasileiro de Radiologia
e Diagnóstico por Imagem

Certificado

Declaramos que o trabalho **Desenvolvimento de rede neural convolucional para o diagnóstico radiográfico de osteoartrite dos joelhos no Estudo Longitudinal da Saúde do Adulto – Musculoesquelético (ELSA-Brasil MSK)** – sob autoria de Júlio Guerra Domingues, Daniella Castro Araújo, Luciana Costa-Silva, Alexei Manso Corrêa Machado, Luciana Andrade Carneiro Machado, Adriano Alonso Veloso, Sandhi Maria Barreto, Rosa Weiss Telles, foi apresentado em formato **Painel eletrônico top de área (Apresentação Oral)** no 51º Congresso Brasileiro de Radiologia e Diagnóstico por Imagem (CBR22).

Florianópolis, 3 de setembro de 2022.

Para verificar a autenticidade deste certificado, basta acessar o link validacertificados.iweventos.com.br e usar o código: c1c6cd1642


Dr. Valdair Francisco Muglia
Presidente do CBR


Dra. Luciana Costa
Diretora Científica do CBR


Dr. Ronaldo Hueb Baroni
Diretor Científico Adjunto do CBR

Certificado de premiação: 1º melhor trabalho na área de Inteligência Artificial, Inovação e Telerradiologia



CBR22
51º CONGRESSO BRASILEIRO
DE RADIOLOGIA E
DIAGNÓSTICO POR IMAGEM

cbr | Colégio Brasileiro de Radiologia
e Diagnóstico por Imagem

Declaramos que o trabalho "**Desenvolvimento de rede neural convolucional para o diagnóstico radiográfico de osteoartrite dos joelhos no Estudo Longitudinal da Saúde do Adulto – Musculoesquelético (ELSA-Brasil MSK)**" – sob autoria de Júlio Guerra Domingues, Daniella Castro Araújo, Luciana Costa-Silva, Alexei Manso Corrêa Machado, Luciana Andrade Carneiro Machado, Adriano Alonso Veloso, Sandhi Maria Barreto, Rosa Weiss Telles, foi selecionado como **1º melhor trabalho na área de Inteligência Artificial, Inovação e Telerradiologia**, no 51º Congresso Brasileiro de Radiologia e Diagnóstico por Imagem (CBR22).

Florianópolis, 3 de setembro de 2022.

Para verificar a autenticidade deste certificado, basta acessar o link validacertificados.iweventos.com.br e usar o código: cd810c50f5


Dr. Valdair Francisco Muglia
Presidente do CBR


Dra. Luciana Costa
Diretora Científica do CBR


Dr. Ronaldo Hueb Baroni
Diretor Científico Adjunto do CBR

Certificado de distinção (Magna Cum Laude): 1º melhor trabalho geral



CBR22
51º CONGRESSO BRASILEIRO
DE RADIOLOGIA E
DIAGNÓSTICO POR IMAGEM

 **cbr** | Colégio Brasileiro de Radiologia
e Diagnóstico por Imagem

Certificado

Declaramos que o trabalho "**Desenvolvimento de rede neural convolucional para o diagnóstico radiográfico de osteoartrite dos joelhos no Estudo Longitudinal da Saúde do Adulto – Musculoesquelético (ELSA-Brasil MSK)**" – sob autoria de Júlio Guerra Domingues, Daniella Castro Araújo, Luciana Costa-Silva, Alexei Manso Corrêa Machado, Luciana Andrade Carneiro Machado, Adriano Alonso Veloso, Sandhi Maria Barreto, Rosa Weiss Telles, recebeu a distinção de **MAGNA CUM LAUDE** por ser o 1º melhor trabalho do 51º Congresso Brasileiro de Radiologia e Diagnóstico por Imagem (CBR22).

Florianópolis, 3 de setembro de 2022.

Para verificar a autenticidade deste certificado, basta acessar o link validacertificados.iweventos.com.br e usar o código: 5fdf31e34d


Dr. Valdair Francisco Muglia
Presidente do CBR


Dra. Luciana Costa
Diretora Científica do CBR


Dr. Ronaldo Hueb Baroni
Diretor Científico Adjunto do CBR

II. Trabalho apresentado no British Institute of Radiology Annual Congress 2022

Certificado de apresentação de pôster eletrônico



III. Folha de Aprovação do Comitê de Ética



UNIVERSIDADE FEDERAL DE MINAS GERAIS
COMITÊ DE ÉTICA EM PESQUISA - COEP

Projeto: CAAE 0186.1.203.000-06

Interessado(a): Profa. Sandhi Maria Barreto
Depto. de Medicina Preventiva e Social
Faculdade de Medicina -UFMG

DECISÃO

O Comitê de Ética em Pesquisa da UFMG – COEP analisou e aprovou, no dia 03 de setembro de 2012, a inclusão de exames, abaixo relacionados, na 2ª. etapa de obtenção dos dados (Onda 2) do projeto de pesquisa intitulado “ELSA - Estudo Longitudinal da Saúde do Adulto”:

- Avaliação da força isométrica nas mãos (*handgrip*);
- Teste de força isométrica das pernas (assentar/levantar repetido);
- Teste de sensibilidade com monofilamento;
- Medida de altura abdominal (diâmetro sagital abdominal);
- Avaliação radiológica das mãos e joelhos;
- Circunferência da cabeça;
- Altura do joelho.
- Termo de Consentimento Livre e Esclarecido.

A aprovação é válida por 1(um) ano (03 de setembro de 2012 a 02 de setembro de 2013).

Profa. Maria Teresa Maques Amaral
Coordenadora do COEP/UFMG