

MESTRADO INTEGRADO EM MEDICINA
GASTROENTEROLOGIA

A Inteligência Artificial na Doença Inflamatória Intestinal: Do Diagnóstico à Medicina de Precisão

Eduardo Miguel Lameirinhas Teixeira

M

2026



(página intencionalmente deixada em branco)

A INTELIGÊNCIA ARTIFICIAL NA DOENÇA INFLAMATÓRIA INTESTINAL: DO DIAGNÓSTICO À MEDICINA DE PRECISÃO

Artigo de revisão bibliográfica

Dissertação de candidatura ao grau de Mestre em Medicina, submetida ao Instituto de Ciências
Biomédicas Abel Salazar, Universidade do Porto

O Estudante:

Eduardo Miguel Lameirinhas Teixeira

Estudante do 6.º ano do Mestrado Integrado em Medicina

Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto

Endereço eletrónico: up202008230@up.pt

A Orientadora:

Dr.ª Marta Salgado Rodrigues

Assistente Hospitalar Graduada de Gastrenterologia, Unidade Local de Saúde de Santo António

Professora Auxiliar Convidada da Unidade Curricular de Medicina I e Especialidades Médicas, do
Mestrado Integrado em Medicina do Instituto de Ciências Biomédicas Abel Salazar, Universidade do
Porto

O Coorientador:

Prof. Doutor Fernando Manuel de Castro Poças

Assistente Hospitalar Graduado Sénior de Gastrenterologia, Unidade Local de Saúde de Santo António

Professor Catedrático Convidado com Agregação e Regente da Unidade Curricular de Medicina I e
Especialidades Médicas, do Mestrado Integrado em Medicina do Instituto de Ciências Biomédicas
Abel Salazar, Universidade do Porto

Porto, junho de 2026.

(página intencionalmente deixada em branco)

DECLARAÇÃO DE INTEGRIDADE CIENTÍFICA

Eu, Eduardo Miguel Lameirinhas Teixeira, portador do número mecanográfico 202008230 e estudante do 6.º ano do Mestrado Integrado em Medicina do Instituto de Ciências Biomédicas Abel Salazar, na Universidade do Porto, declaro que esta dissertação é da minha autoria e que não foi utilizada previamente de modo total ou parcial noutro curso ou unidade curricular desta ou de outra instituição.

As referências a outros autores (afirmações, ideias e pensamentos) encontram-se devidamente indicadas no texto e nas referências bibliográficas, de acordo com as regras de referenciação.

Tenho conhecimento de que a prática de plágio e auto-plágio constitui uma infração académica.

Porto, junho de 2026.

(Eduardo Miguel Lameirinhas Teixeira)

Assinatura do Estudante: Eduardo Miguel Lameirinhas Teixeira

Assinatura da Orientadora: Dr.^a Marta Salgado Rodrigues

Assinatura do Coorientador: Prof. Doutor Fernando Manuel de Castro Poças

“A tecnologia é um servo útil, mas um mestre perigoso.”

— Christian Louis Lange (1921)

(página intencionalmente deixada em branco)

AGRADECIMENTOS

À minha orientadora, a Dr.^a Marta Salgado, pela disponibilidade imediata que demonstrou para orientar esta dissertação. Um agradecimento especial pela sugestão deste tema, cujo interesse me cativou desde o primeiro instante; pela leitura atenta e o tempo generoso dedicado a este trabalho, que apenas se tornou possível pela amabilidade com que sempre garantiu o apoio necessário em cada momento.

Ao meu coorientador, o Prof. Doutor Castro Poças, pelo acolhimento imediato e pela orientação nesta área. A sua visão e direcionamento foram fundamentais para a concretização desta dissertação.

A ambos, acima de tudo pelo brio, retidão e simpatia com que sempre desempenharam as suas funções como médicos e, particularmente, como docentes, sendo referências indeléveis no percurso de qualquer aluno desta casa que tenha o privilégio de com ambos contactar.

À minha namorada, que está a trilhar este mesmo caminho e que percebe tão bem as exigências desta etapa. Mesmo estando mais longe do que gostaríamos, obrigado pelo companheirismo, pelos momentos de desabafo e pela motivação quando precisei.

À minha família, que me acompanhou nesta jornada, e aos amigos com quem a partilhei e que lhe deram mais cor.

Por fim, mas nunca em último, aos meus pais, a quem nunca poderei retribuir tudo o que me deram. Obrigado por estarem sempre lá, pelo carinho, pelo amor, pelas lições, pela vossa paciência infinita, por acreditarem em mim mesmo quando eu duvidei e nunca me deixarem baixar os braços. Por se terem colocado tantas vezes em segundo plano para que eu pudesse estar em primeiro. Se estou a terminar esta etapa é porque nunca me faltou o vosso apoio. Esta caminhada é tanto minha quanto vossa.

A todos, o meu obrigado.

(página intencionalmente deixada em branco)

RESUMO

Introdução:

As Doenças Inflamatórias Intestinais (DII), que incluem a Doença de Crohn (DC) e a Colite Ulcerosa (CU), são patologias crônicas que afetam o trato gastrointestinal (TGI) e cujo curso clínico tem impacto na qualidade de vida dos doentes e nos custos em saúde, pelo que o diagnóstico atempado e estratégias personalizadas de tratamento e monitorização são essenciais. Atualmente, estas baseiam-se na avaliação endoscópica, imagiológica e histopatológica, que se associam a subjetividade, variabilidade interobservador e custos elevados. Ademais, a personalização das estratégias de seguimento e terapêutica poderão ainda ser aprimoradas. Neste contexto, e considerando a crescente prevalência destas patologias na comunidade, a Inteligência Artificial (IA), através de abordagens como *Machine Learning* (ML), *Deep Learning* (DL), Radiômica e *Natural Language Processing* (NLP), surge como uma tecnologia capaz de integrar dados complexos, apoiando a decisão clínica.

Objetivos:

Esta revisão narrativa tem como objetivo avaliar o estado da arte da aplicação de IA à DII, particularmente na acuidade do diagnóstico diferencial e caracterização da doença, bem como o seu potencial na investigação e na medicina de precisão, especialmente no seguimento, previsão do curso da doença e da resposta terapêutica. Procura-se ainda examinar as suas limitações atuais, questões éticas suscitadas e perspectivas futuras na atividade clínica.

Metodologia:

Foi realizada uma revisão bibliográfica com recurso ao motor de pesquisa *PubMed*, recorrendo inicialmente à combinação de palavras-chave “((Artificial Intelligence) OR (Machine Learning) OR (Deep Learning)) AND ((Inflammatory Bowel Disease) OR (Crohn* Disease) OR (Ulcerative Colitis))”.

A seleção de artigos foi feita tendo em conta, sequencialmente, o título, resumo e corpo, tendo sido valorizados os estudos mais recentes que apresentassem métricas de desempenho diagnóstico ou explorassem a integração da IA na prática clínica. A bibliografia dos artigos inicialmente incluídos foi também analisada para identificar outros artigos relevantes não contemplados na pesquisa inicial. Estas fontes foram complementadas por outros recursos quando apropriado, nomeadamente de sociedades científicas e consensos clínicos.

Discussão:

A IA demonstra aplicação transversal na abordagem à DII. No diagnóstico diferencial, permite diferenciar DII de outras patologias e entre DC e CU com base na avaliação de dados clínicos,

endoscópicos e histológicos. Ainda baseada nestas modalidades, revela-se útil na caracterização da doença ao automatizar a aplicação de escalas endoscópicas e histológicas, prever remissão histológica com base em dados endoscópicos e, na DC, reduzir o tempo de análise de videocápsula endoscópica (VCE) sem comprometer a acuidade diagnóstica. Através de modalidades imagiológicas, a IA auxilia na distinção fenotípica e atividade da doença de forma não invasiva, evidenciando a importância da multimodalidade neste domínio.

No contexto da medicina de precisão, os modelos de IA apresentam potencial na previsão do curso da doença e da resposta terapêutica, com evidência crescente de que ao incorporarem informação longitudinal no tempo têm maior poder preditivo face aos baseados apenas em dados iniciais.

No seguimento do doente, as plataformas digitais e os *wearables* permitem uma vigilância contínua e não invasiva, revelando-se capazes de antecipar agudizações, permitindo uma atuação mais atempada. Por sua vez, modelos de NLP mostram-se oportunos na educação em saúde, embora exibindo limitações que exigem supervisão médica.

Na investigação, esta tecnologia revela aplicabilidade no recrutamento de doentes, na padronização da avaliação de meios complementares de diagnóstico e terapêutica (MCDT) e no desenvolvimento e reposicionamento de fármacos.

Contudo, permanecem limitações como a heterogeneidade dos dados, a escassa validação multicêntrica, a opacidade algorítmica e questões ético-legais que deverão ser ultrapassadas para permitir a aplicação desta tecnologia na prática médica corrente.

Conclusão:

A evidência sintetizada mostra que a IA apresenta uma trajetória promissora na abordagem da DII, ao permitir maior acuidade e objetividade no diagnóstico e caracterização da doença. Ademais, demonstra particular potencial no seguimento do doente, na previsão do curso da doença e da resposta à terapêutica, influenciando positivamente a qualidade de vida. Porém, a sua translação clínica depende do estabelecimento de bases ético-legais sustentadas e que as suas limitações atuais sejam ultrapassadas através da validação em populações heterogêneas e da criação de modelos explicáveis que integrem dados multimodais e longitudinais, o que contribuirá para a maior personalização dos cuidados de saúde. Em suma, a IA deve ser entendida como uma ferramenta de apoio à decisão médica, potenciando o juízo clínico sem abdicar da insubstituível dimensão humana do médico.

Palavras-Chave: Inteligência Artificial; Doença Inflamatória Intestinal; Doença de Crohn; Colite Ulcerosa

ABSTRACT

Introduction:

Inflammatory bowel diseases (IBD), which include Crohn's disease (CD) and ulcerative colitis (UC), are chronic conditions that affect the gastrointestinal tract (GIT) and whose clinical course impacts patients' quality of life and healthcare costs, making timely diagnosis and personalized treatment and monitoring strategies essential. Currently, these strategies rely on endoscopic, imaging, and histopathological evaluations, which are associated with subjectivity, inter-observer variability, and high costs. Furthermore, the personalization of follow-up and therapeutic strategies could be further improved. In this context, and considering the growing prevalence of these conditions in the community, Artificial Intelligence (AI), through approaches such as Machine Learning (ML), Deep Learning (DL), Radiomics, and Natural Language Processing (NLP), emerges as a technology capable of integrating complex data to support clinical decision-making.

Objectives:

This narrative review aims to assess the current state of the art in the application of AI to IBD, particularly with regard to the accuracy of the differential diagnosis and disease characterization, as well as its potential in research and precision medicine, particularly in monitoring, predicting disease progression, and therapeutic response. It also seeks to examine its current limitations, the ethical issues it raises, and future prospects in clinical practice.

Methods:

A literature review was conducted using the PubMed search engine, initially employing the keyword combination “((Artificial Intelligence) OR (Machine Learning) OR (Deep Learning)) AND ((Inflammatory Bowel Disease) OR (Crohn* Disease) OR (Ulcerative Colitis))”.

Articles were selected by sequentially reviewing their title, abstract, and main text, with priority given to the most recent studies that presented diagnostic performance metrics or explored the integration of AI into clinical practice. The references of the initially included articles were also analyzed to identify other relevant articles not included in the initial search. These sources were supplemented by other resources when appropriate, notably from scientific societies and clinical consensus statements.

Discussion:

AI has broad applications in the management of IBD. In differential diagnosis, it enables the differentiation of IBD from other conditions and between CD and UC based on the evaluation of clinical,

endoscopic, and histological data. Based on these modalities, it proves useful in characterizing the disease by automating the application of endoscopic and histological scales, predicting histological remission based on endoscopic data, and, in CD, reducing the time required for video capsule endoscopy analysis without compromising diagnostic accuracy. Through imaging modalities, AI assists in the non-invasive distinction of phenotypes and disease activity, highlighting the importance of multimodality in this field.

In the context of precision medicine, AI models show potential in predicting disease course and therapeutic response, with growing evidence that by incorporating longitudinal information over time, they have greater predictive power compared to those based solely on initial data.

In patient follow-up, digital platforms and wearables enable continuous, non-invasive monitoring, proving capable of anticipating exacerbations and allowing for more timely intervention. In turn, NLP models are proving useful in health education, though they exhibit limitations that require medical supervision.

In research, this technology has proven useful in patient recruitment, in standardizing the evaluation of ancillary studies, and in drug development and repositioning.

However, limitations remain, such as data heterogeneity, limited multicenter validation, algorithmic opacity, and ethical and legal issues that must be overcome to enable the application of this technology in routine medical practice.

Conclusion:

The synthesized evidence shows that AI holds promise in the management of IBD, as it enables greater accuracy and objectivity in the diagnosis and characterization of the disease. Furthermore, it demonstrates particular potential in patient follow-up, in predicting disease course and response to therapy, and in positively influencing quality of life. However, its clinical translation depends on the establishment of sound ethical and legal foundations and on overcoming its current limitations through validation in heterogeneous populations and the creation of interpretable models that integrate multimodal and longitudinal data, which will contribute to greater personalization of healthcare. In short, AI should be understood as a tool to support medical decision-making, enhancing clinical judgment without sacrificing the irreplaceable human dimension of the physician.

Keywords: Artificial Intelligence; Inflammatory Bowel Disease; Crohn's Disease; Ulcerative Colitis

LISTA DE ABREVIATURAS

AGA — *American Gastroenterological Association*
AINE — Anti-Inflamatório Não Esteroide
AUC — *Area Under the Curve*
BWT — Espessura da Parede
CADe — Detecção Assistida por Computador
CADx — Diagnóstico Assistido por Computador
CCR — Cancro Colorretal
CDEIS — *Crohn's Disease Index of Severity*
CDS — *Cumulative Disease Score*
CECDAl — *Capsule Endoscopy Crohn's Disease Activity Index*
CI — Colite Isquêmica
CNN — *Convolution Neural Networks*
CONSORT-AI — *Consolidated Standards of Reporting Trials—Artificial Intelligence*
CPF — Calprotectina Fecal
CU — Colite Ulcerosa
CXCR4 — *C-X-C Chemokine Receptor Type 4*
DBI — Doença de Behçet Intestinal
DC — Doença de Crohn
DII — Doença Inflamatória Intestinal
DL — *Deep Learning*
EnteroRM — Enterografia por Ressonância Magnética
EnteroTC — Enterografia por Tomografia Computadorizada
ESGE — Sociedade Europeia de Endoscopia Gastrointestinal
FDA — *Food and Drug Administration*
GEDII — Grupo de Estudos da Doença Inflamatória Intestinal
GHAS — *Global Histological Activity Score*
GS — *Geboes Score*
IA — Inteligência Artificial
ICC — *Intraclass correlation coefficient*
IL — Interleucina
JAK — *Janus Kinase*
LLM — *Large Language Models*
LS — *Lewis Score*
MaRIA — *Magnetic Resonance Index of Activity*
MCDT — Meios Complementares de Diagnóstico e Terapêutica
MES — *Mayo Endoscopic Score*

ML — *Machine Learning*
NHI — *Nancy Histological Index*
NLP — *Natural Language Processing*
pCLE — *Endomicroscopia Laser Confocal*
PCR — *Proteína C Reativa*
PHRI — *PICaSSO Histologic Remission Index*
PKC θ — *Protein Kinase θ*
QWK — *Quadratic Weighted Kappa*
RD — *Red Density*
RHI — *Robarts Histopathology Index*
S1P — *Esfingosina-1-Fosfato*
SES-CD — *Simple Endoscopic Score for Crohn's Disease*
SHAP — *SHapley Additive exPlanations*
SII — *Síndrome de Intestino Irritável*
SPIRIT-AI — *Standard Protocol Items: Recommendations for Interventional Trials–Artificial Intelligence*
STRIDE II — *Selecting Therapeutic Targets in Inflammatory Bowel Disease II*
SWE — *Single Wavelength Endoscopy*
TBI — *Tuberculose Intestinal*
TGI — *Trato Gastrointestinal*
TNF α — *Tumor Necrosis Factor α*
UC-SCALE — *Ulcerative Colitis Severity Classification and Localized Extent*
UCEIS — *Ulcerative Colitis Endoscopic Index of Severity*
UST — *Ustekinumab*
VCE — *Videocápsula Endoscópica*
VDZ — *Vedolizumab*
VPN — *Valor Preditivo Negativo*
VPP — *Valor Preditivo Positivo*

ÍNDICE

Agradecimentos	i
Resumo	iii
Abstract	v
Lista de Abreviaturas	vii
Índice	ix
Lista de Tabelas	xi
Lista de Figuras.....	xii
Introdução.....	1
Objetivos.....	2
Métodos.....	2
1. Inteligência Artificial no Diagnóstico Diferencial	5
1.1. Diagnóstico Diferencial entre Doença Inflamatória Intestinal e Outras Patologias	5
1.2. Diagnóstico Diferencial entre Colite Ulcerosa e Doença de Crohn	6
2. Inteligência Artificial na Caracterização da Doença	7
2.1. Endoscopia e Atividade Endoscópica	7
2.1.1. Aplicação na Colite Ulcerosa	8
2.1.2. Aplicação na Doença de Crohn.....	9
2.1.2.1. Redução dos Tempos de Leitura	10
2.1.2.2. Caracterização de Lesões	10
2.1.2.3. Plataformas Comerciais	11
2.1.3. Previsão da Atividade Histológica	11
2.2. Histopatologia e Atividade Histológica	13
2.3. Aplicação em Imagiologia	14
2.3.1. Ecografia	14
2.3.2. Enterografia por Ressonância Magnética	14
2.3.3. Enterografia por Tomografia Computadorizada	15
3. Inteligência Artificial em Medicina de Precisão	16
3.1. Previsão do Curso da Doença	16
3.1.1. Previsão de Agudizações	16
3.1.2. Previsão de Complicações	17
3.1.3. Previsão da Necessidade de Cirurgia	17
3.1.4. Previsão de Complicações Pós-Operatórias.....	18

3.2. Seleção de Fármacos e Previsão da Resposta à Terapêutica	18
3.2.1. Apoio à Decisão Terapêutica	19
3.2.2. Resposta à Terapêutica Convencional	19
3.2.3. Resposta a Anti-TNF α	20
3.2.4. Resposta ao Vedolizumab	20
3.2.5. Resposta ao Ustekinumab	21
3.2.6. Resposta Terapêutica em Ensaios Clínicos	21
4. Inteligência Artificial no Seguimento do Doente	22
4.1. Vigilância e Detecção de Neoplasias	22
4.1.1. Papel da Inteligência Artificial nas Técnicas Endoscópicas	22
4.1.2. Papel da Inteligência Artificial na Histologia	23
4.2. Monitorização e Educação do Doente.....	23
4.2.1. Plataformas de Saúde Digital.....	24
4.2.2. Dispositivos e <i>Wearables</i>	25
4.2.3. Educação do Doente.....	25
5. Inteligência Artificial em Investigação.....	26
5.1. Otimização de Ensaios Clínicos.....	26
5.2. Desenvolvimento Farmacológico	27
6. Limitações Atuais, Questões Éticas e Perspetivas Futuras.....	27
Conclusão	29
Apêndice.....	31
Bibliografia	39

LISTA DE TABELAS

Tabela I — Tabela Comparativa Entre as Principais Características de DC e CU	31
Tabela II — Tabela Comparativa Entre CADe e CADx	32
Tabela III — Tabela Comparativa Entre Biomarcadores Tradicionais e Digitais na DII	33

LISTA DE FIGURAS

Figura 1 — Esquematização dos Principais Subtipos de IA	34
Figura 2 — Processo de Desenvolvimento e Avaliação de um Modelo de IA Hipotético	36
Figura 3 — Esquematização do Processo de Recolha e Análise Bibliográfica	38

INTRODUÇÃO

As DII são patologias crônicas, que afetam o TGI, e incluem a DC e a CU que apresentam características clínicas, endoscópicas, histopatológicas, genéticas e epidemiológicas distintas.¹ A DC pode afetar qualquer porção do TGI, associando-se a inflamação transmural e complicações como estenoses, fístulas e abscessos.² Na CU ocorre inflamação contínua da mucosa, limitada ao cólon e reto (Tabela I).^{1,2} Ademais, o curso clínico destas patologias tem impacto direto na qualidade de vida do doente e nos custos associados a hospitalizações, cirurgias e terapêuticas avançadas, tornando essencial um diagnóstico correto e atempado, bem como estratégias personalizadas de tratamento e monitorização.³

A avaliação do doente sustenta-se na avaliação de MCDT de modalidades como a endoscopia, imagiologia e histopatologia, cuja interpretação tende a ser subjetiva, sujeita a vieses, onerosa e morosa.^{3,4} Adicionalmente, embora a intervenção terapêutica deva ser personalizada às comorbilidades, tratamentos prévios, manifestações extraintestinais e preferências individuais do doente, o seu correto enquadramento em todos estes domínios permanece uma tarefa difícil.^{3,4}

Atualmente, estima-se que a prevalência da DII ronde os 1% nos países desenvolvidos,^{1,2,5} mas globalmente verifica-se uma tendência crescente, que poderá sobrecarregar os cuidados de saúde, pelo que se tem reforçado a procura de novas estratégias capazes de agilizar o diagnóstico e potenciar a decisão clínica.⁵

Assim, a IA surge como uma tecnologia capaz de integrar conjuntos de dados vastos e complexos para apoiar a gestão da doença. Esta refere-se à aplicação de sistemas computacionais para resolver problemas que tipicamente requerem inteligência humana e, presentemente, abrange vários modelos aplicados a múltiplos domínios.⁶⁻⁸

Como sugerido pela Figura 1, a ML é um subconjunto de IA definido por algoritmos capazes de analisar dados para identificar padrões e de completar tarefas sem serem especificamente programados para tal. Estes são geralmente treinados de modo supervisionado a partir de dados rotulados.^{6,8}

O DL insere-se na ML e caracteriza-se por uma arquitetura algorítmica de várias camadas, permitindo o ajuste dos parâmetros de cada uma, otimizando os *outputs* e possibilitando a abordagem de tarefas mais complexas⁶. As *Convolution Neural Networks* (CNN) são a arquitetura mais característica que, inspiradas na organização do córtex visual, são principalmente aplicadas na identificação de padrões visuais em imagens, útil para interpretação endoscópica, histológica e imagiológica (Figura 2).^{6,8} Ainda

aplicada à imagiologia, a Radiômica permite extrair características quantitativas de imagens médicas, auxiliando a sua avaliação.^{6,9}

Outra técnica é a NLP, que permite extrair informação a partir de texto não estruturado, como notas clínicas ou relatórios de MCDT, possibilitando depois a sua análise mais rigorosa. Os *Large Language Models* (LLM), como o ChatGPT™, enquadram-se neste domínio.^{10,11}

Assim, a IA terá implicações no diagnóstico, caracterização, tratamento e monitorização da DII, ao recorrer a modelos treinados na interpretação de dados digitalizados, como endoscopias, biópsias, imagens médicas ou processos clínicos, facilitando a identificação de padrões dificilmente detetáveis por humanos.⁸

As suas principais aplicações na DII associam-se à interpretação rápida e uniforme de MCDT, colmatando as lacunas supramencionadas, e à categorização de doentes quanto à evolução da doença e resposta à terapêutica, viabilizando estratégias individualizadas para melhorar os desfechos clínicos.^{8,12}

OBJETIVOS

A presente dissertação tem como objetivo investigar o estado da arte da aplicação de IA na DII, através de uma revisão crítica da literatura existente. Pretende-se avaliar o impacto desta tecnologia na acuidade e objetividade do diagnóstico diferencial e caracterização da doença, bem como o seu potencial na investigação e na medicina de precisão, particularmente no seguimento, previsão do curso da doença e da resposta terapêutica. Procura-se ainda examinar as suas limitações atuais, questões éticas suscitadas e perspetivas futuras na atividade clínica.

MÉTODOS

Esta dissertação assenta numa revisão narrativa efetuada com recurso ao motor de pesquisa *PubMed*, recorrendo inicialmente à combinação de palavras-chave “((Artificial Intelligence) OR (Machine Learning) OR (Deep Learning)) AND ((Inflammatory Bowel Disease) OR (Crohn* Disease) OR (Ulcerative Colitis))”.

Os artigos foram selecionados (Figura 3) tendo em conta, sequencialmente, o título, resumo e corpo do artigo, sendo excluídos quando duplicados ou se abordassem temas fora do âmbito desta revisão. Foram incluídos e lidos na totalidade os artigos que atendiam aos seguintes critérios: artigos originais,

de revisão, ou meta-análises, redigidos em língua inglesa, com pertinência para os objetivos da dissertação e com disponibilidade do artigo completo *online*.

Foram valorizados os estudos mais recentes que apresentassem métricas de desempenho diagnóstico ou explorassem a integração da IA na prática clínica. A bibliografia dos artigos incluídos foi também analisada para identificar outros artigos relevantes não contemplados na pesquisa inicial. Estas fontes foram complementadas por outros recursos sempre que considerado pertinente, nomeadamente por documentos de sociedades científicas e consensos clínicos.

A aplicação *Zotero*[™] foi utilizada para gerir a literatura.

(página intencionalmente deixada em branco)

1. INTELIGÊNCIA ARTIFICIAL NO DIAGNÓSTICO DIFERENCIAL

1.1. DIAGNÓSTICO DIFERENCIAL ENTRE DOENÇA INFLAMATÓRIA INTESTINAL E OUTRAS PATOLOGIAS

O diagnóstico diferencial das DII engloba patologias distintas com características clínicas, endoscópicas e histológicas que se sobrepõem, dificultando a sua discriminação.¹³ A IA, ao auxiliar na interpretação de MCDT, poderá aumentar a acuidade desta diferenciação.¹³

Em países desenvolvidos, alguns dos principais diferenciais de DII são a colite isquêmica (CI) e a colite infecciosa.¹⁴ Contudo, a análise isolada de imagens endoscópicas pode ser subótima nesta distinção, como ilustrado por Guimarães et al., cujo modelo baseado em parâmetros clínicos teve desempenho superior na diferenciação destas patologias comparativamente ao baseado na análise de imagens endoscópicas e até mesmo a endoscopistas experientes, reforçando a importância da multimodalidade nesta avaliação.¹⁴

Nesta direção, têm sido desenvolvidos sistemas capazes de diferenciar várias patologias. O modelo de Luo et al. extrai o contexto espacial e temporal a partir de vídeos de colonoscopia e foi capaz de distinguir entre cólon saudável, DC, CU, CI e tuberculose intestinal (TBI) com exatidão comparável a peritos (80,8% vs. 79,7%) e superior à de endoscopistas menos experientes (71,8%).¹⁵ Noutro estudo, um algoritmo baseado em imagens endoscópicas e parâmetros clínicos revelou maior exatidão que endoscopistas seniores na distinção entre cinco patologias, incluindo doença de Behçet intestinal (DBI) e linfoma intestinal, validando esta abordagem.¹⁶

Também a distinção entre DC e TBI, embora menos prevalente em países desenvolvidos, permanece um desafio significativo, particularmente em contextos com recursos limitados.¹³ A par de algoritmos endoscópicos capazes de diferenciar entre DII, DBI e TBI, modelos que incorporam variáveis clínicas e radiológicas otimizaram a acuidade diagnóstica e contribuíram para a interpretabilidade do resultado ao aplicarem o método *SHapley Additive exPlanations* (SHAP).¹⁷

Por sua vez, a Radiômica aplicada à Enterografia por Tomografia Computadorizada (EnteroTC) demonstrou que características imagiológicas da parede intestinal, do mesentério e dos gânglios linfáticos, particularmente quando combinadas com dados clínicos e endoscópicos, superaram consistentemente o julgamento de radiologistas experientes na distinção entre DC e TBI, reforçando a utilidade desta modalidade diagnóstica.¹⁸

O microbioma é outra área investigada neste contexto. Este complexo ecossistema de microorganismos que coexistem no TGI pode alterar-se em caso de doença, refletindo perturbações fisiopatológicas do hospedeiro e funcionando como uma possível assinatura patológica.¹⁹ Neste sentido, modelos de IA têm sido capazes de distinguir doentes com DII de indivíduos saudáveis e distinguir DII de patologias funcionais, como a Síndrome de Intestino Irritável (SII), sublinhando o papel da disbiose na fisiopatologia destas doenças.²⁰

1.2. DIAGNÓSTICO DIFERENCIAL ENTRE COLITE ULCEROSA E DOENÇA DE CROHN

A distinção entre CU e DC é essencial para a correta abordagem clínica dos doentes, já que a terapêutica e o prognóstico são distintos.^{1,2} Todavia, esta diferenciação não é simples, devido à sobreposição de características clínicas, imagiológicas e histológicas, pelo que a IA tem sido aplicada na análise de dados de múltiplos domínios diagnósticos.^{3,4}

Na avaliação endoscópica, a utilização de modelos de IA tem permitido diferenciar imagens de mucosa normal, DC e CU com acuidade global superior à de endoscopistas e em menos tempo que estes.²¹ Paralelamente, tem potenciado técnicas endoscópicas mais avançadas, como a endomicroscopia laser confocal (pCLE), capaz de uma ampliação de cerca de 1000 vezes. A aplicação desta técnica à análise da arquitetura da mucosa (distância intercripta, espessura da parede e extravasamento de fluoresceína) permitiu distinguir entre indivíduos saudáveis e com DII com sensibilidade e especificidade de 100% e entre doentes com DC e CU com 92% de sensibilidade e 91% de especificidade.²²

Contudo, ao demonstrar que um modelo baseado em dados endoscópicos e histológicos foi superior ao de cada modalidade isoladamente, Mossotto et al. reafirmaram a importância da multimodalidade.²³

Assim, na histologia, Rymarczyk et al. validaram um modelo na distinção da presença de DC ou CU no cólon e de DC no íleo, alcançando boa concordância com o determinado por patologistas.²⁴ Outro algoritmo, capaz de quantificar automaticamente a plasmocitose basal, demonstrou que a prevalência deste grupo celular foi um fator diferenciador entre CU, DC e controlos saudáveis (55,16 vs. 38,22 vs. 17,25).²⁵ Mais recentemente, Cai et al. apresentaram um modelo capaz de distinguir entre doentes saudáveis e DII e entre DC e CU com *Area Under the Curve* (AUC) superiores a 0,95. Notavelmente, apresentou desempenho diagnóstico superior ao de todos os patologistas e, ao assisti-los, melhorou

significativamente a sua *performance*, sublinhando o potencial enquanto ferramenta de auxílio à decisão médica.²⁶

No domínio imagiológico, a EnteroTC suplementada por IA tem permitido avaliar características estruturais como adiposidade visceral e gordura mesentérica, cujos padrões, por serem distintos entre DC e CU, permitiram desenvolver sistemas capazes de diferenciar estas patologias muito satisfatoriamente, particularmente quando esta determinação incluiu também características clínicas.²⁷

Complementarmente, assiste-se à utilização crescente de outras abordagens não invasivas para guiar o diagnóstico. Neste contexto, parâmetros laboratoriais serológicos, urinários e fecais obtidos por rotina demonstraram capacidade para diagnosticar DII e diferenciar entre DC e CU com excelente acuidade diagnóstica.²⁸

Ademais, embora analisar a biologia molecular e genética do doente não seja prática corrente, a literatura atual sustenta o seu potencial futuro. Efetivamente, um dos primeiros artigos demonstrou que dados genéticos permitem distinguir entre indivíduos saudáveis e DC (AUC=0,86) ou CU (AUC=0,83).²⁹ Recentemente, um modelo baseado num painel de genes associados à autoimunidade conseguiu diferenciar DC e CU razoavelmente (AUC=0,68), demonstrando, particularmente, a importância do gene *NOD2* nesta discriminação.³⁰ A análise do microbioma é igualmente promissora no diagnóstico de DII (AUC=0,992) e na distinção entre DC e CU (AUC=0,988).³¹

Finalmente, técnicas de NLP têm permitido extrair informação com valor preditivo diretamente de registos clínicos e relatórios endoscópicos com boa capacidade discriminativa entre DC e CU, ilustrando a sua utilidade na análise de informação amplamente disponível.³²

2. INTELIGÊNCIA ARTIFICIAL NA CARACTERIZAÇÃO DA DOENÇA

2.1. ENDOSCOPIA E ATIVIDADE ENDOSCÓPICA

A endoscopia desempenha um papel central na caracterização da DII, permitindo avaliar a inflamação da mucosa e confirmar a sua cicatrização, alvo terapêutico com importante valor prognóstico.³³ Esta avaliação é objetivada através de escalas padronizadas, como o *Mayo Endoscopic Score* (MES) ou *Ulcerative Colitis Endoscopic Index of Severity* (UCEIS) na CU e o *Simple Endoscopic Score for Crohn's Disease* (SES-CD) ou *Crohn's Disease Index of Severity* (CDEIS) na DC,^{1,2,4} mas a classificação manual é um processo moroso e subjetivo, limitações que poderão ser ultrapassadas pelo recurso à IA.⁸

Neste domínio, o objetivo final da aplicação de IA tem sido o desenvolvimento de sistemas automatizados de detecção e caracterização de lesões (*Computer-Aided Detection* [CAdE], *Computer-Aided Diagnosis* [CAdx] — Tabela II).^{34,35} Todavia, pelas diferenças fenotípicas entre as doenças, as abordagens estudadas têm divergido para responder aos desafios endoscópicos particulares de cada uma.^{1,2,34}

2.1.1. APLICAÇÃO NA COLITE ULCEROSA

A colonoscopia permanece como o exame de referência na avaliação da atividade da CU.² O estudo da aplicabilidade da IA nesta área iniciou-se pela análise de imagens estáticas, para distinguir remissão endoscópica de doença ativa. Um dos primeiros modelos avaliados, por Stidham et al., permitiu distinguir “remissão” (MES 0-1) de “doença moderada-severa” (MES 2-3) com robustez (AUC=0,966) e boa concordância com humanos (κ =0,84), similar à descrita entre endoscopistas (κ =0,86).³⁶ Subsequentemente, Takenaka et al. apresentaram um modelo treinado em 40758 imagens de 2012 doentes, cuja exatidão (90,1%) e concordância interobservador (κ =0,798) corroboraram a aplicabilidade da IA neste domínio. Notavelmente, conseguiu prever a remissão histológica analisando exclusivamente imagens endoscópicas, sugerindo uma potencial aplicação na redução da necessidade de biópsias e a capacidade destes modelos preverem informação sobre uma modalidade distinta da inicialmente avaliada.³⁷

Ainda que promissores, estes modelos analisaram imagens estáticas, acarretando viés de seleção ao não aferirem segmentos de análise potencialmente mais difícil. Para mitigar esta limitação, estudos mais recentes focam-se na análise de vídeo em tempo-real. Neste contexto, Gottlieb et al. aplicaram um modelo à análise de colonoscopias realizadas no ensaio de fase 2 do mirikizumab. Este determinou a remissão endoscópica, definida por UCEIS e MES, com excelente exatidão (97% vs. 95,5%) e concordância (*Quadratic Weighted Kappa* [QWK]=0,885 vs. QWK=0,844), sugerindo a aplicabilidade destes sistemas na investigação, ao auxiliarem na leitura centralizada de MCDT, com potencial para reduzir o tempo e os custos despendidos neste processo.³⁸ Adicionalmente, a transição para a análise de vídeo permitiu novos avanços, como a exclusão de *frames* de má qualidade e a geração de mapas topográficos bidimensionais que ilustram a carga inflamatória ao longo de todo o cólon.^{39,40}

Ultrapassada a fase de avaliação binária (remissão vs. atividade), o passo seguinte consistiu em treinar modelos na graduação da doença em todos os seus níveis de severidade, com o grupo de Qi et al. a apresentar um algoritmo que, com base em imagens endoscópicas, previu a classificação MES com boa exatidão (90,8%) e AUC (>0,973).⁴¹ Esta evolução culminou no desenvolvimento do primeiro sistema

totalmente automatizado, capaz de processar vídeos de colonoscopia e fornecer classificações contínuas de MES e UCEIS com excelente concordância com peritos.⁴²

Finalmente, a literatura sugere que a competência dos algoritmos depende das escalas com as quais são treinados, sendo que o uso de índices mais detalhados, como o UCEIS, produz resultados globalmente mais robustos do que outros mais simples, como o MES.⁴³

Contudo, estas escalas tendem a valorizar somente o segmento cólico mais afetado, o que, embora necessário para que operadores humanos as consigam determinar, poderá ser limitante para modelos de IA. Assim, têm emergido novas métricas desenhadas especificamente para serem calculadas por estas sistemas e que, embora sejam inviáveis para uso humano, poderão potenciar estes algoritmos. Assim, Stidham et al. propuseram o *Cumulative Disease Score* (CDS) que, ao quantificar a atividade inflamatória cumulativa ao longo de todo o cólon, através da avaliação contínua do MES, provou ser tão sensível a pequenas melhorias na mucosa que, quando aplicada aos resultados do ensaio UNIFI, permitiria reduzir o tamanho da amostra em 50%, com óbvias implicações no desenho de futuros ensaios clínicos.⁴⁴

Baseado num método similar, Gutierrez Becker et al. propuseram a *Ulcerative Colitis Severity Classification and Localized Extent* (UC-SCALE) que demonstrou boa concordância com endoscopistas ($\kappa=0,80$) e correlação com a avaliação clínica ($r=0,45$) e biomarcadores de atividade da doença como a Calprotectina Fecal (CPF; $r=0,50$) e Proteína C Reativa (PCR; $r=0,45$), sublinhando a validade desta abordagem.⁴⁵

2.1.2. APLICAÇÃO NA DOENÇA DE CROHN

Na DC, a natureza descontínua e transmural da inflamação e o frequente atingimento do intestino delgado têm desfavorecido a aplicação da IA à endoscopia convencional,^{1,3} tendo o foco da investigação recaído na VCE, enquanto método não invasivo que permite a visualização de toda a mucosa entérica. Contudo, a sua utilização na prática clínica é limitada pelo tempo necessário para a sua avaliação e pela variabilidade interobservador.⁴⁶ Efetivamente, um único exame gera dezenas de milhares de imagens captadas ao longo de várias horas de gravação, pelo que a sua leitura exige atenção sustentada, potenciando o erro humano e comprometendo a deteção de lesões.^{1,46}

2.1.2.1. REDUÇÃO DOS TEMPOS DE LEITURA

Um dos impactos mais imediatos da aplicação desta tecnologia tem-se verificado na drástica redução dos tempos de leitura através de modelos que filtram imagens redundantes e sinalizam *frames* com maior probabilidade de evidenciarem lesões.

Assim, Aoki et al. avaliaram a IA aplicada à VCE como um primeiro filtro de triagem seguido de revisão humana. Este algoritmo permitiu a redução do tempo de leitura por endoscopistas experientes (de 12,2 para 3,1 minutos) e inexperientes (de 20,7 para 5,2 minutos), sem reduzir significativamente a taxa de deteção de lesões em cada grupo. Contudo, os gastroenterologistas inexperientes, embora auxiliados por esta ferramenta, falharam a deteção de lesões significativas, ilustrando que a experiência do leitor condiciona a eficácia do sistema.⁴⁷ Adicionalmente, a aplicação de um algoritmo à remoção de *frames* com má visualização reduziu o tempo de leitura em 35,6% sem perda de informação diagnóstica.⁴⁸

Embora auspiciosos, estes artigos não foram especificamente dirigidos à DC, lacuna explorada por Brodersen et al. que, recorrendo à plataforma AXARO®, conseguiram reduzir o tempo de leitura de VCE pan-entérica para uma mediana de 3,2 minutos (comparativamente a 303 minutos de gravação) mantendo boa sensibilidade e especificidade (>90%) para o diagnóstico de DC. Por sua vez, o elevado valor preditivo negativo (VPN=95%) revela o potencial desta abordagem no rastreio e exclusão rápida desta patologia.⁴⁹

2.1.2.2. CARACTERIZAÇÃO DE LESÕES

Para além da deteção de lesões, a investigação tem-se focado na sua caracterização. Aoki et al. foram pioneiros ao apresentarem um modelo capaz de detetar úlceras e erosões em imagens de VCE de intestino delgado (AUC=0,958). Contudo, não foi aplicado especificamente na DC, lacuna mitigada por outro grupo, cujos algoritmos detetaram úlceras (AUC=0,99),⁵⁰ estenoses (AUC=0,989) e diferenciaram-nas entre si (AUC=0,971),⁵¹ mas revelaram limitações na discriminação entre úlceras de DC e induzidas por Anti-Inflamatórios Não Esteroides (AINE).⁵²

Para além de diferenciar lesões, foi demonstrada a capacidade de a IA graduar automaticamente as úlceras de acordo com a escala de classificação da PillCam™. O sistema avaliado por Barash et al. distinguiu satisfatoriamente entre graus 1 (ligeiro) e 3 (severo) (AUC=0,958) e graus 2 (moderado) e 3 (AUC=0,939), mas com menor robustez entre graus 1 e 2 (AUC=0,565), ilustrando as limitações na distinção de graus intermédios de severidade. Notavelmente, a concordância entre clínicos (31%) foi inferior à demonstrada pelo algoritmo (67%), sugerindo a reprodutibilidade desta técnica.⁵³

Tal como na CU, esta capacidade de deteção e classificação automatizada das lesões na DC potencia o desenvolvimento de escalas padronizadas, dado que o cálculo manual de métricas validadas para VCE como o *Lewis Score* (LS) ou o *Capsule Endoscopy Crohn's Disease Activity Index* (CECDAI) é moroso.⁵⁴ Esta vertente foi explorada por Cardoso et al., que desenvolveram uma escala baseada na percentagem de *frames* identificados pelo algoritmo como apresentando lesões. Esta demonstrou boa correlação com as já validadas, fornecendo uma alternativa automatizada à avaliação manual.⁵⁵

2.1.2.3. PLATAFORMAS COMERCIAIS

A maturidade da aplicação de IA à VCE reflete-se na disponibilização comercial de diversas plataformas. Atualmente, o dispositivo mais utilizado mundialmente na DC é o PillCam™ Crohn (Medtronic®), que permite a avaliação panentérica da doença. Embora o *software* de leitura inclua algumas funcionalidades de leitura assistida, potencialmente úteis na DC,⁵⁶ não possui quaisquer algoritmos específicos para esta patologia, lacuna que tem sido explorada por terceiros, resultando no desenvolvimento da plataforma AXARO® e de outros algoritmos ainda em investigação.^{49,54}

Um dos grupos mais proeminentes nesta área é nacional, tendo já apresentado um modelo robusto na deteção de lesões na DC.⁵⁷ Subsequentemente, estes resultados foram validados num estudo multicêntrico e multidispositivo, passo essencial para a adoção clínica generalizada.⁵⁸

Outro grupo português avaliou a plataforma OMOM® SmartScan Plus™, demonstrando que a triagem por IA com revisão humana pode substituir a revisão completa do vídeo, reduzindo o tempo de leitura sem comprometer a classificação do LS e CECDAI.⁵⁹

Também recentemente, foi apresentado um modelo capaz de detetar lesões (AUC>0,98; exatidão=99,33%) e diagnosticar DC (AUC>0,98; exatidão=90%) similarmente a especialistas, mas 10 vezes mais rápido e, simultaneamente, melhorar o desempenho dos clínicos, particularmente dos menos experientes.⁶⁰

Estes resultados ilustram que a integração da IA na análise da VCE ultrapassou já o plano teórico, estabelecendo-se hoje como uma realidade cada vez mais presente.⁵⁴

2.1.3. PREVISÃO DA ATIVIDADE HISTOLÓGICA

Embora a atividade histológica apresente estreita relação com o prognóstico da DII, a sua avaliação depende da realização de biópsias, cujo carácter invasivo limita a sua aplicação clínica.³³ Porém, como descrito anteriormente, alguns modelos foram capazes de inferir a atividade histológica a partir de dados endoscópicos, impulsionando a investigação nesta vertente.³⁷

Efetivamente, um dos primeiros modelos a demonstrar esta aplicabilidade foi apresentado por Takenaka et al. Baseado em imagens de colonoscopia convencional, previu a remissão histológica na CU com notável exatidão (92,9%) e concordância com a avaliação efetuada por patologistas humanos ($\kappa=0,859$).³⁷ Mais recentemente, novos modelos confirmaram estes resultados, ao superarem a avaliação visual da mucosa por gastroenterologistas na previsão de remissão histológica, mantendo excelente concordância com a avaliação histopatológica ($\kappa=0,826$).⁶¹

Na pesquisa por maior acuidade, a investigação avançou para a análise colorimétrica e para o uso de espectros de luz limitados. Neste contexto, o método de *Red Density* (RD) quantifica automaticamente a inflamação baseando-se na densidade da cor vermelha da mucosa. Este correlaciona-se com escalas endoscópicas estabelecidas como MES ($r=0,76$) e UCEIS ($r=0,74$) e permitiu, ainda, determinar a atividade histológica medida pelo *Robarts Histological Index* (RHI) (AUC=0,95).⁶² Paralelamente, a tecnologia de *Single Wavelength Endoscopy* (SWE) associada à IA alcançou uma exatidão na previsão da atividade histológica (86%) superior à avaliação humana baseada nas escalas de MES (74%) e UCEIS (79%).⁶³ A comparação direta da SWE com a colonoscopia tradicional na previsão da atividade histológica demonstrou superioridade desta nova técnica quanto à sensibilidade (88.0% vs. 73.9%), especificidade (71,7% vs. 65,6%) e exatidão (83,3% vs. 67,5%), provando a utilidade da exploração de diferentes comprimentos de onda de luz na caracterização da doença.⁶⁴

Uma aproximação ainda mais real à histologia *in vivo* foi alcançada através de técnicas de ultramagnificação. A aplicação de IA à endocitoscopia (técnica que permite magnificação até 520 vezes, possibilitando a avaliação da arquitetura críptica, infiltração celular e alterações microvasculares) permitiu prever a atividade histológica persistente (*Geboes Score* [GS] ≥ 3) com boa sensibilidade (74%), especificidade (97%) e exatidão (91%), garantindo ainda uma reprodutibilidade algorítmica perfeita ($\kappa=1$).⁶⁵ Adicionalmente, verificou-se que a avaliação por endocitoscopia apresenta correlação mais forte com índices histológicos ($r>0,85$) do que endoscópicos (MES, $r=0,28$), sublinhando que esta técnica está mais próxima da histologia do que da endoscopia convencional.⁶⁶

Numa abordagem similar, a pCLE assistida por IA alcançou excelente sensibilidade (94,6%), especificidade (92,8%) e exatidão global (95,3%) na distinção entre inflamação ativa e remissão na DC, cimentando o potencial destas modalidades menos invasivas na redução do número de biópsias necessárias, sem comprometer a personalização do seguimento do doente.⁶⁷

2.2. HISTOPATOLOGIA E ATIVIDADE HISTOLÓGICA

A “remissão profunda”, caracterizada pela cicatrização e remissão histológica, é um alvo terapêutico emergente na DII, demonstrando superioridade face à remissão endoscópica isolada na predição do risco de recidiva, necessidade de hospitalização, uso prolongado de corticoides e desenvolvimento de complicações a longo prazo.^{3,4,33} Contudo, a aplicação de índices validados, como o RHI, o GS ou o *Nancy Histological Index* (NHI), é limitada pela variabilidade entre patologistas, lacuna que a IA poderá colmatar.⁸

Na DC, a natureza descontínua e transmural da inflamação impossibilita que uma dada biópsia capte fielmente a atividade em todo o TGI.^{1,4} A maior complexidade do infiltrado inflamatório e a ausência de escalas de classificação formalmente validadas nesta doença são outras limitações que justificam a investigação ser centrada preponderantemente na CU.¹

Num dos primeiros artigos focados neste domínio, Vande Casteele et al. apresentaram um modelo direcionado para a identificação e quantificação de eosinófilos em biópsias do cólon.⁶⁸ Embora tenha produzido contagens celulares com elevada concordância face à contagem manual (*Intraclass correlation coefficient* [ICC]=0,805-0,917), não revelou correlação com a atividade histológica, corroborando o consenso de que o infiltrado neutrofílico tem significado patológico superior, redirecionando os estudos subsequentes nessa direção.⁶⁸

Consequentemente, foi desenvolvido o *PICaSSO Histologic Remission Index* (PHRI), um índice histológico assente exclusivamente na presença ou ausência de neutrófilos, que demonstrou boa correlação com a atividade endoscópica e desfechos clínicos. A sua praticidade foi objetivada pela elevada concordância entre patologistas (ICC=0,84). Um sistema desenvolvido para automatizar o PHRI distinguiu CU ativa de quiescente (exatidão=86%), validando a sua aplicabilidade.⁶⁹

Para além destas novas métricas desenvolvidas para aplicação de IA, outros estudos têm abordado a automatização de escalas histológicas pré-existentes. Modelos capazes de segmentar tecidos e identificar tipos celulares conseguiram determinar o NHI com fiabilidade comparável à humana (κ =0,91) e avaliaram a remissão histológica com exatidão de 97%.⁷⁰ Notavelmente, o algoritmo desenhado por Peyrin-Biroulet et al. para determinar o NHI demonstrou concordância com a avaliação por patologistas (ICC=87,2) praticamente sobreponível à concordância entre patologistas especialistas (ICC=89,3), sugerindo que esta tecnologia terá atingido um nível de fiabilidade adequado para a implementação clínica.⁷¹

Por fim, embora a DC apresente as dificuldades de avaliação supracitadas, a automatização da sua leitura foi alvo de estudo recente por Rymarczyk et al, que treinaram modelos na determinação do GS na CU e do *Global Histological Activity Score* (GHAS) na DC. Estes alcançaram exatidões de 65-85% na avaliação da CU, de 80-89% na avaliação da DC no cólon e de 65-82% na avaliação da DC no íleo. O desempenho na distinção entre graus de severidade apresentou concordância moderada a substancial relativamente ao estabelecido por um painel de patologistas, provando a aplicabilidade deste método a diferentes segmentos do TGI.²⁴

2.3. APLICAÇÃO EM IMAGIOLOGIA

As técnicas imagiológicas, como ecografia, enterografia por ressonância magnética (EnteroRM) e EnteroTC, permitem a caracterização e monitorização não invasivas da DII, com particular aplicabilidade na DC, dada a natureza transmural da inflamação e o frequente atingimento do intestino delgado, inacessível à endoscopia convencional.³³ Embora existam índices para quantificar a atividade da doença e o dano estrutural, como o *Magnetic Resonance Index of Activity* (MaRIA), a sua aplicação clínica é limitada pelo tempo despendido e pela variabilidade interobservador, lacunas que a aplicação da IA procura colmatar.^{4,72}

2.3.1. ECOGRAFIA

Embora a ecografia seja uma ferramenta acessível e não invasiva, a sua principal fragilidade é depender da experiência do operador.^{4,72} Aqui, a espessura da parede (BWT) surge como um bom parâmetro na avaliação da atividade da DC, pelo que tem sido um dos principais focos da investigação.^{8,72}

Procurando mitigar a subjetividade da avaliação humana, Carter et al. automatizaram esta medição através de um modelo capaz de detetar espessamento patológico (BWT>3mm) com boa sensibilidade (86,4%), especificidade (94,0%), exatidão (90,1%) e AUC (0,98).⁷³

Posteriormente, esta abordagem foi validada na população pediátrica, alcançando elevada concordância com gastroenterologistas (ICC=0,942).⁷⁴ Adicionalmente, Gu et al. evidenciaram que modelos Radiómicos baseados na análise da textura intestinal superiorizaram-se a CNN clássicas na avaliação de BWT e hiperémia intestinal (AUC=0,98 vs. AUC=0,75), captando com maior detalhe a anatomia inflamatória.⁷⁵

2.3.2. ENTEROGRAFIA POR RESSONÂNCIA MAGNÉTICA

Por sua vez, a aplicação de IA a imagens obtidas por EnteroRM tem permitido avanços na avaliação anatómica e funcional. Estruturalmente, têm sido apresentados algoritmos capazes de segmentar

automaticamente o lúmen e as camadas da parede intestinal, tornando estas medições significativamente mais reprodutíveis do que se efetuadas manualmente ($ICC=0,88$ vs. $ICC=0,45$).⁷⁶

Ferramentas semiautomáticas recentes permitiram identificar e segmentar lesões, cujo volume se relacionou com a atividade radiológica da doença, e avaliar a gordura mesentérica, cuja hipertrofia se associou mais à remodelação crónica do que à inflamação aguda.⁷⁷

Paralelamente, modelos de Radiómica demonstraram, face à leitura por radiologistas, superioridade na identificação de fibrose grave ($AUC=0,78$ vs. $AUC=0,35$). Notavelmente, embora, isoladamente, a capacidade de identificação de inflamação severa seja idêntica entre radiologistas e algoritmos ($AUC=0,67$), o seu uso pelos primeiros permitiu incrementar substancialmente a acuidade diagnóstica ($AUC=0,79$), ilustrando o papel de suplementação, e não suplantação, da IA neste domínio.⁷⁸

Funcionalmente, a quantificação automática da motilidade intestinal, frequentemente alterada devido à inflamação ou fibrose, revelou-se útil na deteção da DC ($AUC=0,78$).⁷⁹ A variação temporal deste parâmetro e a área de intestino móvel associaram-se à severidade dos sintomas.⁸⁰

Quanto à classificação da atividade da doença, modelos Radiómicos provaram ter um desempenho idêntico ao cálculo manual do MaRIA ($AUC=0,87$ vs. $AUC=0,88$), mas com melhor reprodutibilidade ($ICC=0,93-0,96$ vs. $ICC=0,58$).⁸¹ Em populações mais vulneráveis, como a pediátrica, a análise combinada de marcadores bioquímicos e imagens de EnteroRM estimou a atividade endoscópica com maior acuidade comparativamente ao MaRIA isoladamente, revelando um potencial futuro onde a monitorização rigorosa da DC seja cada vez menos invasiva.⁸²

2.3.3. ENTEROGRAFIA POR TOMOGRAFIA COMPUTADORIZADA

No domínio da EnteroTC, a investigação tem incidido principalmente na distinção fenotípica e na avaliação de danos estruturais na DC. Relativamente à fibrose intestinal, Yuan et al. validaram um modelo baseado nas características da adiposidade visceral e da estenose na determinação da sua severidade.⁸³

Já Li et al. apresentaram um sistema capaz de distinguir fibrose “moderada/severa” de “ligeira/ausente” ($AUC=0,89$) com desempenho superior ao de radiologistas ($AUC=0,55-0,60$).⁸⁴ Ainda neste domínio, a comparação direta entre diferentes arquiteturas computacionais revelou que, embora modelos de DL e Radiómica evidenciem desempenho diagnóstico similar na classificação da fibrose, a abordagem de DL é significativamente mais rápida (48,4s vs. 599,8s), vantagem que pode ter impacto na escolha do algoritmo a utilizar na prática clínica.⁸⁵

Quanto à inflamação, a integração de características Radiômicas do lúmen e do mesentério revelou-se útil para mapear a doença por segmentos, prevendo de modo não invasivo a atividade endoscópica da DC e identificando os segmentos intestinais com doença mais severa.⁹ Outro grupo demonstrou, ainda, que modelos de IA aplicados a este domínio são capazes de elevada acuidade na distinção entre DC clinicamente ativa e em remissão, consolidando o papel crescente destas modalidades imagiológicas na monitorização do doente.⁸⁶

3. INTELIGÊNCIA ARTIFICIAL EM MEDICINA DE PRECISÃO

3.1. PREVISÃO DO CURSO DA DOENÇA

A história natural da DII é heterogênea, variando desde formas indolentes até trajetórias rapidamente progressivas e incapacitantes.^{4,12} A frequência e gravidade das agudizações, a resposta a uma dada terapêutica e o desenvolvimento de complicações são eventos maioritariamente imprevisíveis aquando do diagnóstico.^{1-3,12}

Historicamente, a abordagem terapêutica tem sido reativa ao agravamento clínico, resultando frequentemente em atrasos e falências terapêuticas.^{3,4} Assim, o objetivo atual é inverter este paradigma, transitando de uma medicina reativa para uma abordagem preditiva, onde a IA pode potenciar a antecipação do curso da doença.¹²

3.1.1. PREVISÃO DE AGUDIZAÇÕES

O primeiro passo na alteração do curso da DII passa por antecipar precocemente o seu agravamento. Na CU, tem-se destacado a análise histológica assistida por modelos de IA. No estudo de Iacucci et al., o grupo identificado como histologicamente ativo pelo modelo baseado no PHRI evidenciou risco relativo de 4,64 de recorrência a 12 meses.⁸⁷

Igualmente, a quantificação automática de depleção de células caliciformes na mucosa associou-se a um risco de recidiva de 45% a 12 meses, contrastando com apenas 6,5% nos não depletados.⁸⁸ Complementarmente, a avaliação funcional da cicatrização da barreira epitelial, através do estudo da integridade das *tight-junctions*, com recurso a pCLE, endocitoscopia e imunofluorescência, demonstrou que a densidade celular e a expressão de Claudina-2 se relacionaram com os desfechos adversos a 12 meses, cimentando o papel da barreira mucosa na previsão de agudizações.⁸⁹

Para além da microscopia, a aplicação da IA a dados clínicos e laboratoriais de rotina afigura-se como uma direção de investigação pragmática, pela ubiquidade da avaliação de biomarcadores como a CPF neste contexto. Efetivamente, a aplicação de um modelo baseado na CPF e complementado com

informações demográficas e do histórico da doença permitiu melhorar a previsão da recidiva na CU comparativamente ao uso isolado deste biomarcador (AUC=0,754 vs. AUC=0,619).⁹⁰

Adicionalmente, a extração de dados dos registos eletrónicos por NLP e a sua avaliação através de modelos de DL têm-se demonstrado promissores na previsão de agudizações a curto prazo, embora a sua generalização ainda careça de validação clínica robusta.⁹¹

3.1.2. PREVISÃO DE COMPLICAÇÕES

Se o processo inflamatório crónico não for atempadamente frenado, espera-se a progressão da doença, com o desenvolvimento de complicações, como fístulas e estenoses na DC.¹ Nesta patologia, a integração de variáveis clínicas, caracterização genética do gene *NOD2* e painéis de proteínas séricas já demonstraram a capacidade de prever complicações em populações adultas e pediátricas.^{92,93}

Já Kurowski et al. aplicaram um modelo de NLP aos dados clínicos de doentes com DC capaz de a classificar quanto à sintomatologia e à atividade com 85% e 96% de exatidão, respetivamente. Dos doentes classificados como assintomáticos, mas com doença ativa, mais de 25% evoluíram para sintomáticos em 1 ano, corroborando a importância e valor prognóstico dos registos clínicos.⁹⁴

3.1.3. PREVISÃO DA NECESSIDADE DE CIRURGIA

Quando as complicações se instalam ou a terapêutica falha, a cirurgia torna-se uma opção, com importantes consequências para a morbilidade do doente e a gestão dos recursos em saúde.¹⁻³ A previsão da necessidade de cirurgia tem sido avaliada através de duas vias complementares: dados clínico-laboratoriais e imagiologia.

No domínio clínico, vários estudos em doentes com CU estratificaram o risco de proctocolectomia sustentado em parâmetros como os valores de PCR, albumina, a estratégia terapêutica corrente e a classificação endoscópica.^{95,96}

Notavelmente, Magro et al. recorreram à base de dados do Grupo de Estudos da DII (GEDII) para o desenvolvimento de matrizes de risco na CU, determinando que doentes com doença extensa, previamente hospitalizados, sem terapêutica imunomoduladora e refratários à corticoterapia têm risco elevado de realizarem colectomia (88%). O modelo preditivo baseado nestes dados alcançou AUC de 0,92 na previsão da necessidade desta cirurgia.⁹⁷

Na DC, o modelo de Stidham et al. conseguiu estratificar o risco de cirurgia aos 12 meses (AUC=0,78). O uso de corticoides foi identificado como fator de risco e a classe dos anti-*Tumor Necrosis Factor α*

(TNF α) como protetora. Dos marcadores laboratoriais, contagens elevadas de plaquetas, concentrações médias elevadas de hemoglobina celular, níveis baixos de albumina e de azoto ureico sérico foram identificados como fatores de risco para desfechos adversos.⁹⁸

No domínio imagiológico, a aplicação de EnteroTC à obtenção de características Radiômicas relativas à adiposidade visceral permitiu prever a progressão fenotípica e a necessidade de cirurgia na DC.⁹ A aplicação de modelos preditivos às imagens de EnteroRM demonstrou AUC de 0,74 na previsão do risco de cirurgia a 12 meses, bem como boa correlação com o tempo até cirurgia, validando o papel de métodos não-ionizantes neste contexto.⁹⁹

3.1.4. PREVISÃO DE COMPLICAÇÕES PÓS-OPERATÓRIAS

Apesar do seu papel na abordagem à DII, a ressecção cirúrgica, particularmente na DC, não representa a cura, podendo ocorrer complicações pós-operatórias e recorrência da doença.¹

No período pós-operatório precoce, modelos de IA têm permitido prever complicações como a deiscência anastomótica, através da avaliação de critérios de risco pré-operatórios.¹⁰⁰ Notavelmente, Mascarenhas et al. desenvolveram um sistema que, através da avaliação de *frames* de vídeos laparoscópicos da anastomose durante a cirurgia, conseguiu determinar quase perfeitamente a presença de deiscência anastomótica (AUC=0,996; Valor Preditivo Positivo [VPP]=100%; VPN=98,9%). Embora não tenha sido aplicado na DII, a utilização nesta patologia constitui uma promissora via de investigação futura.¹⁰¹

A longo prazo, a previsão da recorrência pós-cirúrgica tem sido avaliada principalmente na DC, onde a avaliação das características clínicas tem permitido o desenvolvimento de modelos preditivos do risco de recorrência.¹⁰² Também a análise histológica tem tido um papel preponderante, como evidenciado pelo modelo de Kiyokawa et al., capaz de prever a recorrência a 2 anos (AUC=0,995) pela redução do volume dos adipócitos e infiltração mastocitária na subserosa.¹⁰³

Já Wang et al. obtiveram resultados mais modestos na validação externa de um modelo baseado na avaliação da inflamação do plexo mioentérico (AUC=0,87), ilustrando a incerteza inerente à previsão do curso da doença.¹⁰⁴

3.2. SELEÇÃO DE FÁRMACOS E PREVISÃO DA RESPOSTA À TERAPÊUTICA

O arsenal terapêutico para abordar a DII tem-se expandido nas últimas décadas com a incorporação de múltiplas classes farmacológicas, que modulam as vias do TNF α , das integrinas, das interleucinas (IL) 12 e 23, da Janus Kinase (JAK) e do recetor Esfingosina-1-Fosfato (S1P).³

Embora traduzida na melhoria dos desfechos clínicos dos doentes, esta expansão tornou a decisão clínica num processo altamente complexo, porque a seleção, sequenciação e combinação farmacológica adequada se baseia frequentemente em consensos científicos e na escalada terapêutica após falência das abordagens anteriores, numa lógica de "tentativa e erro".^{1-3,105} Esta estratégia pouco individualizada pode associar-se a maiores custos e ao atraso no controlo da doença, com mais de 30% dos doentes a não responderem inicialmente ou a perderem resposta ao longo do tempo.^{4,8}

Neste contexto, a IA tem o potencial para personalizar a abordagem ao doente, nomeadamente quanto à escolha farmacológica ideal e à previsão da resposta à terapêutica selecionada.¹²

3.2.1. APOIO À DECISÃO TERAPÊUTICA

Um dos primeiros desafios clínicos consiste na identificação atempada dos doentes que necessitam de escalar para terapêutica biológica. Modelos baseados em informação clínica facilmente obtida, como a idade do doente, medicação utilizada e valores de ferro sérico, conseguiram prever esta necessidade com 74% de exatidão.¹⁰⁶

Por sua vez, um algoritmo aplicado à análise de VCE de doentes recentemente diagnosticados com DC e ainda não sujeitos a tratamento previu a necessidade de escalada para terapêutica biológica (AUC=0,86), superando o desempenho de marcadores clássicos, como a CPF (AUC=0,74), e do LS determinado por gastroenterologistas (AUC=0,70).¹⁰⁷

Uma vez estabelecida a necessidade de terapêutica biológica, a IA pode também auxiliar na seleção da classe farmacológica a iniciar. Ferramentas de apoio à decisão clínica permitiram estratificar doentes e antecipar se estes beneficiariam mais de opções com ação intestinal seletiva, como o vedolizumab (VDZ), ou sistémica, como anti-TNF α ou ustekinumab (UST), demonstrando benefício superior à escolha empírica tradicional.^{108,109} Paralelamente, abordagens multiómicas com base na avaliação de perfis metabólicos e imunológicos conseguem classificar o fenótipo de cada doente, orientando a escolha terapêutica para anti-TNF α ou VDZ.¹¹⁰

3.2.2. RESPOSTA À TERAPÊUTICA CONVENCIONAL

Para além da escolha do fármaco, um dos pilares da medicina de precisão é a previsão da eficácia da terapêutica instituída, que se pode aplicar aos fármacos clássicos e biológicos. Aliás, um dos primeiros trabalhos nesta área foi elaborado por Waljee et al. e demonstrou que modelos baseados em dados laboratoriais e na idade dos doentes superou a análise clássica de metabolitos de 6-tioguanina na previsão da remissão clínica em doentes submetidos a terapêutica imunomoduladora com tiopurinas

(AUC=0,79 vs. AUC=0,49).¹¹¹ Contudo, uma análise da literatura alerta para que, apesar do seu valor preditivo, estes modelos ainda não possuem a robustez necessária para substituir completamente a monitorização convencional na prática clínica.¹¹²

Quanto aos aminossalicilatos, em doentes com CU ligeira a moderada, algoritmos baseados em dados clínico-laboratoriais simples como o rácio linfócitos/PCR e CPF, bem como na avaliação endoscópica e histológica, demonstraram excelente capacidade para prever a resposta à messalazina.^{113,114}

3.2.3. RESPOSTA A ANTI-TNF α

Relativamente às terapêuticas anti-TNF α , como o infliximab e adalimumab, os estudos publicados têm sublinhado a importância da avaliação longitudinal da doença. Neste aspeto, Con et al. desenvolveram um modelo para prever a remissão da DC. Quando este apenas incluiu dados do início do tratamento, teve um desempenho similar ao modelo de regressão estatística clássica (AUC=0,710 vs. AUC=0,659, $p=0,25$), mas, ao incluir dados do seguimento do doente, revelou maior poder discriminativo (AUC=0,754, $p=0,036$).¹¹⁵

Outros artigos têm corroborado a importância da avaliação da variação dos valores laboratoriais ao longo do tempo, tendo a evidência recaído sobre a hemoglobina, plaquetas, velocidade de sedimentação, PCR, CPF, albumina, níveis de complemento, gordura subcutânea e estado nutricional.^{116,117} A avaliação destes parâmetros bioquímicos pode ser complementada pela análise de imagens obtidas por EnteroTC e EnteroRM, com resultados encorajadores neste domínio.^{118,119}

A nível molecular, foi demonstrado que a expressão aumentada de *DPY19L3* e *GSTT1* e diminuída de *NUCB1* se correlacionavam com a falência terapêutica de anti-TNF α na DC.¹²⁰ Ademais, análises multiómicas longitudinais demonstraram que as alterações na expressão génica à segunda semana de tratamento são preditores muito mais robustos da remissão futura do que o perfil genético basal do doente.¹²¹

3.2.4. RESPOSTA AO VEDOLIZUMAB

Quanto à terapêutica anti-integrinas com VDZ, um dos principais estudos avaliou um modelo que se destacou pelo elevado VPN (92,3%), ou seja, pela capacidade de identificar precocemente (até à semana 22) os doentes que não irão responder ao tratamento, poupando tempo e recursos.¹²²

Os resultados de outros estudos têm ainda sublinhado a importância do desenvolvimento de modelos longitudinais, que tenham em conta dados obtidos ao longo da evolução da doença e não apenas no

início do tratamento, e da conjugação entre dados clínicos e laboratoriais para incrementar a acuidade dos modelos preditivos.^{123,124}

Adicionalmente, o perfil metabólico dos linfócitos T reguladores e os enterótipos que compõem o microbioma intestinal de cada indivíduo evidenciaram também uma forte capacidade preditiva da resposta a este fármaco.^{125,126}

3.2.5. RESPOSTA AO USTEKINUMAB

Os estudos focados na avaliação da resposta ao UST têm corroborado as abordagens focadas na evolução temporal dos marcadores, ilustrado pelo estudo de Waljee et al., que demonstrou que a variação dinâmica de biomarcadores séricos como PCR e albumina até à oitava semana de tratamento é melhor preditor da falência terapêutica que os valores basais isolados (AUC=0,78 vs. AUC=0,59). Notavelmente, o rácio albumina/PCR à sexta semana alcançou AUC (0,76) similar à avaliação seriada até à oitava semana, pelo que este poderá ser um preditor de remissão facilmente determinado na prática clínica.¹²⁷

Para além da vertente analítica, a avaliação Radiômica de imagens intestinais obtidas por EnteroTC permitiu prever a eficácia deste fármaco de forma robusta e não invasiva.¹²⁸ Já a integração de dados genómicos evidenciou a importância dos genes *HSD3B1*, *MUC4*, *CF1*, e *CCL11* na previsão da resposta terapêutica (AUC=0,73).¹²⁹

3.2.6. RESPOSTA TERAPÊUTICA EM ENSAIOS CLÍNICOS

Adicionalmente, a IA tem sido aplicada no estudo de fármacos em ensaios clínicos. Efetivamente, os dados do ensaio de fase III do etrolizumab foram analisados por um algoritmo que, recorrendo a SHAP, isolou os principais preditores da remissão na fase de indução (PCR, MAdCAM-1, e frequência de defecação) e na fase de manutenção (leucocitose, CPF e idade), alcançando AUC de 0,75 nesta previsão.¹³⁰

De igual modo, o ensaio de fase II do mirikizumab foi analisado por um modelo híbrido baseado na endoscopia e histologia que determinou a remissão histológica com boa correlação com patologistas ($\kappa=0,78$), prevendo a remissão às 52 semanas com excelente sensibilidade (98%), especificidade (87%) e exatidão (93%), validando a importância da integração de dados multimodais na avaliação deste domínio.¹³¹

Apesar do progresso destes modelos, a escassez de dados associados a classes mais recentes limita o desenvolvimento de modelos preditivos, ilustrando que a recolha e sistematização de dados da prática clínica é fundamental para o desenvolvimento destas tecnologias.¹³²

4. INTELIGÊNCIA ARTIFICIAL NO SEGUIMENTO DO DOENTE

4.1. VIGILÂNCIA E DETECÇÃO DE NEOPLASIAS

O risco de desenvolvimento de cancro colorretal (CCR) em doentes com DII é proporcional à duração, extensão e severidade da doença, pelo que se preconiza o início do rastreio endoscópico a partir do oitavo ano de doença ou no momento de diagnóstico de colangite esclerosante primária, sendo a endoscopia de alta definição e a cromoendoscopia as técnicas de eleição.^{4,105}

Contudo, a deteção precoce de neoplasia nestes doentes é desafiante, pois, ao contrário da população geral, em que os pólipos tendem a ser esporádicos e bem definidos, as lesões displásicas na DII são frequentemente planas, de bordos mal delimitados e surgem numa mucosa já alterada, caracterizada pela inflamação crónica.⁴

4.1.1. PAPEL DA INTELIGÊNCIA ARTIFICIAL NAS TÉCNICAS ENDOSCÓPICAS

Atualmente, estão disponíveis sistemas de CAdE que alertam o operador em tempo real, durante a endoscopia, para a presença de lesões suspeitas.³⁴ Alguns destes também incluem sistemas de CAdx que caracterizam as lesões, diferenciando entre neoplásicas e não-neoplásicas (Tabela II).³⁴

Alguns dos principais sistemas em uso são o GI Genius™, ENDO-AID™, CADEYE™, Discovery™ e EndoBRAIN EYE™, com resultados promissores no auxílio do diagnóstico de CCR na população geral.^{35,133} Este potencial foi confirmado pela recomendação emitida pela Sociedade Europeia de Endoscopia Gastrointestinal (ESGE) a favor da utilização destes sistemas no rastreio de CCR, embora ressalvando que a evidência é ainda limitada, com benefícios reduzidos na mortalidade e possível sobrecarga do doente pelo maior diagnóstico de pólipos e necessidade de vigilância endoscópica, com custos acrescidos, motivos também citados pela Associação Americana de Gastroenterologia (AGA) para não emitir um parecer a favor ou contra o uso de CAdE neste contexto.^{134,135}

Embora promissores, estes sistemas não estão validados na DII, sendo esta patologia explicitamente indicada como critério de exclusão na sua validação, limitando a sua utilização neste contexto.³⁵ Contudo, estudos recentes têm relatado a sua utilidade na DII, com resultados encorajadores na deteção de lesões neoplásicas¹³⁶ e estão reportados casos específicos onde a aplicação destes sistemas permitiu a deteção de displasia em doentes com DII, sugerindo a sua utilidade neste domínio.^{137,138}

Assim, a investigação tem evoluído no sentido de desenvolver ferramentas específicas para a DII. A pertinência desta direção é demonstrada por Vinsard et al. que, inicialmente, aplicaram um modelo treinado em lesões displásicas gerais a doentes com DII, alcançando sensibilidade (50%) e exatidão (64%) limitadas. Após retreinar o algoritmo com lesões específicas de DII, o modelo atingiu excelente sensibilidade (95%), especificidade (99%) e bom AUC (0,85), corroborando a importância da utilização de modelos específicos.¹³⁹

Avançando da deteção para a caracterização das lesões, Yamamoto et al. treinaram um modelo para classificar as lesões detetadas como “displasia de baixo grau/adenoma esporádico/mucosa normal” ou “adenocarcinoma/displasia de alto grau”. Este alcançou sensibilidade de 72,5%, especificidade de 82,9%, e exatidão de 79,0%, superando o desempenho de endoscopistas experientes e não experientes (sensibilidade de 60,5% e 70,5%, especificidade de 88,0% e 78,8%, e exatidão de 77,8% e 75,8%, respetivamente), sublinhando o potencial da IA na uniformização da avaliação endoscópica.¹⁴⁰

4.1.2. PAPEL DA INTELIGÊNCIA ARTIFICIAL NA HISTOLOGIA

Além do papel da IA na endoscopia, este estende-se à confirmação histológica das lesões. Alterações moleculares, como mutações no gene *p53*, ocorrem em fases precoces da carcinogénese associada à colite.¹⁴¹ Assim, Noguchi et al. desenvolveram um modelo capaz de prever os resultados da imunohistoquímica do *p53* através da análise direta de lâminas histológicas coradas com hematoxilina-eosina. Com exatidão entre 86% e 91%, este sistema demonstra-se uma alternativa mais rápida e economicamente mais eficiente face à imunohistoquímica clássica.¹⁴¹

Além da previsão de mutações isoladas, um estudo recente demonstrou a capacidade de classificar lâminas histológicas como correspondendo a cólon saudável, CU ou CCR, alcançando exatidão superior a 99%. Adicionalmente, recorreu a mapas de calor para evidenciar quais as regiões celulares que motivaram o diagnóstico, contribuindo para a explicabilidade do modelo.¹⁴²

Globalmente, estes avanços apontam para o potencial da IA aplicada à endoscopia e histologia no diagnóstico de CCR na DII, posicionando-se como elemento-chave na modernização dos protocolos de vigilância nesta doença.

4.2. MONITORIZAÇÃO E EDUCAÇÃO DO DOENTE

A introdução das *guidelines* STRIDE-II (*Selecting Therapeutic Targets in Inflammatory Bowel Disease II*) cimentou a mudança de paradigma no que respeita a monitorização da DII, ao estabelecer a importância, a curto prazo, da normalização de biomarcadores e do alívio sintomático para que, a longo

prazo, se alcance a remissão clínica, cicatrização endoscópica e restauro da qualidade de vida, sublinhando assim a importância da avaliação continuada do doente.³³

Contudo, historicamente, esta tem sido feita episodicamente, em contexto de consulta hospitalar e dependendo de exames invasivos, que não permitem captar as flutuações diárias da atividade da doença e que sobrecarregam os sistemas de saúde.¹³² Neste contexto, a saúde digital surge como mitigadora destas limitações, possibilitando a monitorização contínua e em domicílio do doente (Tabela III).^{132,143}

4.2.1. PLATAFORMAS DE SAÚDE DIGITAL

Uma das linhas de investigação atuais prende-se com plataformas de saúde digital. Na prática, estas ferramentas correspondem a portais *web* ou aplicações móveis que não só têm recursos educativos para o doente, como permitem a sua monitorização remota e a comunicação com o clínico caso seja necessário.¹⁴³

Uma das primeiras a ser implementada foi a Constant Care™, desenvolvida na Dinamarca em 2009.¹⁴³ Esta consiste num portal *web* no qual os doentes inserem periodicamente os resultados do teste de CPF realizado no domicílio e respondem a questionários relativos à sintomatologia, bem-estar e qualidade de vida.¹⁴³ Os resultados permitem determinar os índices de *Total Inflammation Burden* e o *Wellbeing*, que são depois utilizados pela equipa médica para ajuste de medicação ou investigações adicionais, se apropriado.¹⁴³ Desde a sua criação, esta plataforma tem demonstrado permitir ajustes terapêuticos com impacto positivo no número de consultas realizadas, qualidade de vida e adesão terapêutica, embora sem efeito significativo nas hospitalizações ou necessidade de cirurgia.¹⁴³ Contudo, está disponível apenas na Dinamarca e necessita da realização de testes de CPF em domicílio, limitando a sua aplicabilidade universal.¹⁴³

Mais recentemente, a ubiquidade no acesso a dispositivos móveis como *smartphones* tem levado ao foco no desenvolvimento de aplicações móveis, como TECCU™ (Espanha) ou myIBDcoach™ (Países Baixos).¹⁴³ Relativamente à plataforma espanhola, um estudo demonstrou que os doentes seguidos via TECCU™ tiveram taxas de remissão mais elevadas e um menor número de consultas presenciais, com provável impacto económico positivo.¹⁴⁴ Quanto à plataforma neerlandesa, a predição de agudizações tendo também em conta medidas autorreportadas revelou desempenho superior (AUC=0,77) comparativamente ao modelo que apenas avaliou variáveis clínicas isoladas (AUC=0,65), sublinhando o papel destas ferramentas.¹⁴⁵

Embora os doentes tendam a reportar satisfação com o uso destas plataformas, a adesão a longo prazo pode constituir um desafio, particularmente em doentes em remissão prolongada, sugerindo a importância da personalização conforme a severidade e fase da doença.¹⁴³

4.2.2. DISPOSITIVOS E *WEARABLES*

Além do registo manual de dados pelo doente, tem-se investigado a pertinência da monitorização passiva da DII, impulsionada pela crescente utilização de *wearables* na população geral, por motivos de bem-estar ou lazer, mesmo fora do contexto de doença.¹⁴⁶ Já que estes dispositivos são utilizados continuamente e recolhem dados sem interrupção da rotina do utilizador, são ótimos candidatos a desempenhar esta função.¹⁴⁶

Neste contexto, têm sido desenvolvidos sensores específicos, como o SWEATSENER™ e o IBD AWARE™, integrados em adesivos ou pulseiras e capazes de quantificar biomarcadores no suor, como IL-1β, IL-6, PCR e CPF, com boa correlação com valores séricos e fecais, mas limitados na previsão da atividade da doença.^{147,148}

Outros estudos têm explorado a avaliação de parâmetros fisiológicos como o número de passos diários, a frequência cardíaca em repouso, a variabilidade da frequência cardíaca e a oxigenação.¹⁴⁶ Por serem métricas fisiológicas facilmente avaliadas por dispositivos de utilização corrente como Apple Watch™, Fitbit™, ou Oura Ring™, sustentam a hipótese da utilização destes aparelhos neste contexto.¹⁴⁶ Efetivamente, Hirten et al. demonstraram que alterações nestes parâmetros permitiram prever exacerbações até sete semanas antes da sua ocorrência, bem como diferenciar entre doentes com e sem inflamação ativa, com potencial de atuar como alerta precoce, antecipando o agravamento clínico antes do doente o perceber.¹⁴⁶

4.2.3. EDUCAÇÃO DO DOENTE

Na Medicina em geral e na DII em particular, a literacia em saúde é essencial, desempenhando um papel crucial na adesão terapêutica e na gestão autónoma da doença.¹³² Neste contexto, os LLM, como o ChatGPT™, Claude™ ou Gemini™, têm-se destacado pela fácil e cada vez mais universal utilização no quotidiano.¹⁰

Estas ferramentas revelam potencial no esclarecimento de dúvidas básicas, gerando respostas mais completas e com grau de empatia similar às fornecidas pelos médicos.^{10,11} No entanto, a interpretação das respostas carece de espírito crítico, particularmente para questões sobre temas com maior complexidade, como imunossupressão, vacinação, planeamento familiar e rastreios oncológicos,

porque tendem a desviar-se das *guidelines* mais atuais, realçando a preocupação quanto à capacidade destes modelos se manterem na vanguarda do conhecimento médico.^{10,11}

Adicionalmente, a inconsistência na geração de respostas, tendencialmente distintas para uma mesma questão, e a utilização de referências omissas comprometem a reprodutibilidade e confiança no sistema, pelo que, embora democratizem o acesso à informação, a integração clínica destas ferramentas exigirá a supervisão e validação pelo médico assistente.^{10,11}

5. INTELIGÊNCIA ARTIFICIAL EM INVESTIGAÇÃO

O processo de investigação é, em medicina, fulcral para o desenvolvimento de evidência sólida na qual seja possível sustentar a prática médica.^{8,149} A DII, enquanto patologia heterogênea na sua apresentação, evolução e resposta ao tratamento, é também dependente da evolução do conhecimento para proporcionar aos doentes os melhores cuidados.^{1,2,149}

Contudo, a investigação na DII enfrenta desafios significativos, como dificuldades na triagem de doentes adequados e elevado consumo de recursos e tempo no desenvolvimento de ensaios clínicos e novas moléculas.^{8,149}

5.1. OTIMIZAÇÃO DE ENSAIOS CLÍNICOS

Relativamente à otimização de ensaios clínicos, a IA auxilia desde as fases inaugurais da investigação. Efetivamente, ferramentas como o ChatGPT™ têm sido avaliadas no processo de ideação da hipótese a estudar. Embora demonstrem limitações quanto à originalidade e especificidade das propostas, revelam a capacidade de estruturar e resumir a literatura, sugerindo ideias cientificamente relevantes.¹⁵⁰

O recrutamento é uma das etapas mais difíceis no desenvolvimento de ensaios clínicos, tendo-se vindo a verificar um declínio sustentado nas taxas de recrutamento, com impacto óbvio neste processo.¹⁴⁹

Um dos principais motivos apontados é a falha na identificação de candidatos, com estudos recentes a sugerir que métodos tradicionais identificam menos 38% dos candidatos quando comparados a novas ferramentas com base em IA.¹⁵¹ Assim, a aplicação de modelos de NLP à análise combinada de códigos de codificação de doença e de dados clínicos não estruturados, como diários ou medicação, facilita a identificação de características clínicas, histológicas ou fenotípicas, de forma a selecionar coortes mais homogêneas, reduzindo o tempo despendido neste processo.¹⁵²

Uma vez recrutados os doentes, um dos principais desafios é a necessidade de avaliar um grande volume de MCDT. Portanto a sua automatização contribuirá para reduzir a variabilidade interobservador, garantindo um processo mais célere, padronizado e menos oneroso.^{12,38}

5.2. DESENVOLVIMENTO FARMACOLÓGICO

Um dos objetivos dos ensaios clínicos passa pela avaliação da segurança e eficácia de novos fármacos. Na DII, esta é uma prioridade importante, pois, apesar da expansão das classes farmacológicas e dos agentes que estão hoje ao dispor, mantêm-se limitações na sua eficácia e efeitos adversos.^{3,105} Assim, modelos de IA podem acelerar a identificação de novos potenciais alvos moleculares, principalmente através da análise da expressão genética e proteómica.¹⁴⁹

Outros modelos têm sido usados no desenvolvimento computacional de moléculas *de novo*, como é o caso do Composto XVI, um antagonista do recetor *C-X-C Chemokine Receptor Type 4* (CXCR-4), atualmente em fase pré-clínica, que demonstrou propriedades farmacocinéticas razoáveis, efeitos anti-inflamatórios *in vitro* e *in vivo* e alívio sintomático em modelos animais.¹⁵³ Apesar do entusiasmo em torno desta aplicabilidade da IA, iniciar o desenvolvimento de um fármaco não garante o seu sucesso clínico, já que a maioria acaba por ser cancelada durante este processo, sublinhando a incógnita quanto ao verdadeiro impacto clínico desta técnica.¹⁴⁹

Perante a incerteza, o tempo e os custos elevados na criação de novos fármacos, uma estratégia que tem ganho destaque é o seu reposicionamento, que consiste em identificar fármacos já aprovados e seguros que possam ser eficazes na DII.¹⁵⁴ Um estudo recente aplicou IA à análise da base de dados de efeitos adversos da *Food and Drug Administration* (FDA) para procurar fármacos cujo uso na população geral estivesse associado a um menor risco de reportar eventos de DII. Desta análise foram priorizados nove fármacos, como a lenalidomida, o dupilumab, a ciclofosfamida e o fingolimod, que poderão ser futuramente estudados, abrindo a possibilidade de utilizar dados de farmacovigilância para desenvolver novas opções terapêuticas.¹⁵⁴

6. LIMITAÇÕES ATUAIS, QUESTÕES ÉTICAS E PERSPETIVAS FUTURAS

Apesar do potencial da IA na revolução da abordagem à DII, mantêm-se limitações quanto à sua aplicabilidade. Efetivamente, o predomínio de estudos retrospectivos, unicêntricos e baseados em amostras limitadas eleva o risco de viés e *overfitting* dos modelos, comprometendo a sua generalização.¹⁵⁵ Isto é agravado pela escassez de validação prospetiva e externa, que constitui uma das principais barreiras à translação clínica desta tecnologia.^{132,156} Ademais, a heterogeneidade nos

métodos, dados utilizados e desfechos avaliados limitam a reprodutibilidade e comparação entre estudos.^{72,132}

A nível ético-legal, a opacidade dos algoritmos (fenómeno de “caixa negra”) dificulta a interpretabilidade dos resultados e a compreensão das decisões algorítmicas, debilitando a relação médico-doente e a correção dirigida destes sistemas.¹⁵⁶ Adicionalmente, a sua implementação depende de garantir a privacidade e soberania relativa aos dados médicos.¹⁵⁶ É ainda basilar que a IA se posicione como uma ferramenta de suporte à decisão médica, complementando e não substituindo o juízo clínico e a relação médico-doente, apoiando-se na perceção positiva dos clínicos quanto ao impacto futuro desta tecnologia.^{156,157}

Assim, a investigação futura deverá focar-se na superação destas lacunas. De facto, atualmente assiste-se à evolução do paradigma tecnológico, através de novos métodos como SHAP, que, ao explicarem os resultados, facilitam a sua compreensibilidade e utilização na decisão clínica⁸. Espera-se também que a generalização dos modelos seja facilitada pela aprendizagem federada, multicêntrica e apoiada em protocolos padronizados como SPIRIT-AI (*Standard Protocol Items: Recommendations for Interventional Trials–Artificial Intelligence*) e CONSORT-AI (*Consolidated Standards of Reporting Trials–Artificial Intelligence*).^{72,156} Adicionalmente, a transição de modelos unimodais para a combinação de várias modalidades diagnósticas revela-se um novo paradigma que, combinando dados clínicos, endoscópicos, histológicos e moleculares, permitirá refinar e personalizar o diagnóstico, caracterização e seguimento da DII.¹²

Por fim, existe ainda algum ceticismo médico, particularmente relativo aos custos de implementação, necessidade de formação e o risco da dependência do operador nestes sistemas, pelo que urge colmatar estas lacunas, de modo a contribuir para a adoção desta tecnologia.¹⁵⁷

CONCLUSÃO

Ao longo dos últimos anos, a IA tem emergido como uma das ferramentas com maior potencial transformador na abordagem à DII abrangendo, simultaneamente, várias áreas, desde o diagnóstico diferencial e caracterização da doença, até à investigação e medicina de precisão. A literatura avaliada nesta dissertação demonstra que, independentemente do domínio considerado, os modelos têm evoluído no sentido de igualar ou superar o desempenho humano em tarefas específicas, realizá-las com maior objetividade, recorrer a dados multimodais para obter resultados mais robustos e recorrer a dados de um domínio (endoscópico, por exemplo) para prever informações de outro (como o histológico), ilustrando um futuro em que a abordagem à DII possa ser efetuada com base em MCDT menos invasivos, mitigando o impacto diagnóstico imposto aos doentes.

Desde logo, tem-se demonstrado a crescente capacidade destes modelos para discriminar DC de CU e estas de outras patologias com características sobreponíveis como a TBI, CI e DBI, o que se reveste de importância crescente, particularmente em regiões onde o acesso a médicos experientes é limitado.

Na caracterização da doença, é particularmente notória a evolução diferencial do papel da IA quando aplicada à CU e à DC. Enquanto na CU o foco tem recaído essencialmente no aprimoramento da colonoscopia clássica, na DC a aplicação de IA à VCE destaca-se pela redução dos tempos de leitura sem comprometer a deteção de lesões, encontrando-se cada vez mais integrada em plataformas comerciais com evidência científica crescente. Concomitantemente, na histopatologia e na imagiologia têm sido explorados algoritmos capazes de cumprir tarefas com desempenho idêntico ou superior ao de clínicos experientes.

No âmbito da medicina de precisão, o maior potencial clínico imediato prende-se com a capacidade de prever o curso da doença de cada indivíduo bem como a sua resposta às diferentes classes farmacológicas, com relevância na personalização da abordagem ao doente. Neste contexto, a incorporação da avaliação temporal longitudinal e de dados multimodais afigura-se determinante no desenvolvimento de modelos preditivos robustos.

Quanto ao seguimento do doente, atualmente, a utilização de plataformas digitais e de *wearables* reveste-se de evidência crescente, particularmente ao permitirem uma monitorização contínua e não invasiva, com impacto positivo na adesão terapêutica, antecipação de agudizações e, consequentemente, no bem-estar do doente. Este seguimento é ainda potenciado pela literacia em saúde, onde modelos de LLM como o ChatGPT™ têm vindo a ter influência crescente. Apesar de

associados a maior democratização no acesso à informação, a sua tendência para divergir das recomendações mais atuais reforça a importância da supervisão médica no seu uso e potenciará a propagação de desinformação, pelo que o médico terá um papel crescente na desmistificação de ideias incorretas em saúde e na DII, em particular.

Na investigação, a IA revela competência no recrutamento de doentes, na padronização da avaliação de exames complementares e no desenvolvimento e reposicionamento de fármacos, com potencial para reduzir os custos e o tempo associados à investigação clínica.

Apesar destes avanços, persistem limitações de ordem técnica e ética que limitam a sua translação clínica, como a heterogeneidade dos dados, a escassez de resultados validados em estudos multicêntricos, a opacidade algorítmica e a importância da segurança e soberania dos dados clínicos, questões potencialmente mitigáveis pela adoção de novas arquiteturas explicáveis e de protocolos experimentais padronizados.

Em suma, a IA não pode ser entendida como um substituto do médico, mas sim como qualquer outro MCDT. Assim, o seu valor deve residir na capacidade de sustentar com maior acuidade o julgamento clínico e libertar tempo que possa ser investido na relação médico-doente, contribuindo para uma abordagem à DII mais precisa, atempada e personalizada ao doente, dado que nenhum algoritmo, por mais sofisticado que seja, irá substituir o conhecimento técnico, o sentido crítico e a dimensão humana do médico, características de que dependerá o impacto real desta tecnologia na vida dos doentes.

APÊNDICE

Tabela I — Tabela Comparativa Entre as Principais Características de DC e CU

Baseada em ^{1,2,158}.

	Característica	Doença de Crohn	Colite Ulcerosa
História e Sintomas	Início	Mais abrupto (dias/semanas)	Mais indolente (semanas/meses)
	Localização	Qualquer parte do TGI	Limitada ao cólon (início no reto com progressão proximal)
	Dor abdominal	Muito comum	Pouco comum / Rara
	Hemorragia Retal	Menos frequente	Muito comum
	Tabagismo	Fator de risco	Fator protetor
	Sintomas Constitucionais	Frequentes (febre, perda de peso, fadiga)	Menos frequentes
Exame Físico	Doença Perianal	Muito comum (fistulas, abscessos)	Muito rara
	Dor Abdominal	Comum (frequentemente no quadrante inferior direito)	Menos comum / Difusa
	Manifestações Extraintestinais	Frequentes (olhos, articulações, pele)	Menos frequentes (exceto CEP e Pioderma Gangrenoso)
Achados Laboratoriais	Marcadores Inflamatórios	Elevação marcada de PCR e VS	Elevação ligeira a moderada de PCR e VS
	Anticorpos	ASCA positivo	p-ANCA positivo
	Anemia	Muito comum	Comum
Achados Endoscópicos	Padrão de Distribuição	Descontínuo	Contínuo e simétrico
	Aspetto da Mucosa	Ulcerações aftosas, aspeto em "pedra de calçada", <i>skipping lesions</i>	Eritema difuso, perda do padrão vascular, granularidade, friabilidade, ulcerações superficiais
	Envolvimento do Reto	Frequentemente poupado	Envolvimento quase constante
	Estenoses e Fístulas	Comuns	Muito raras
Achados Histológicos	Profundidade da Inflamação	Transmural (afeta todas as camadas)	Mucosa e submucosa (superficial)
	Granulomas	Caseosos (característicos, mas não patognomônicos)	Ausentes (ou raros)
	Infiltrado Inflamatório	Linfocítico, focal, descontínuo	Linfoplasmocitário difuso, com distorção das criptas e abscessos cripticos

ASCA: anti-Saccharomyces cerevisiae antibodies; **CEP:** Colangite Esclerosante Primária; **p-ANCA:** perinuclear anti-neutrophil cytoplasmic antibodies; **PCR:** Proteína C Reativa; **TGI:** Trato Gastrointestinal; **VS:** Velocidade de Sedimentação.

Tabela II — Tabela Comparativa Entre CAdE e CAdx

Característica	<i>Computer-Aided Detection (CAdE)</i>	<i>Computer-Aided Diagnosis (CAdx)</i>
Objetivo Principal	Detetar e localizar anomalias ou lesões em tempo real durante o exame	Caracterizar, classificar e graduar lesões que já foram detetadas
Pergunta Clínica	<i>"Existe alguma lesão ou área suspeita aqui?"</i>	<i>"Que tipo de lesão é esta e qual a sua gravidade/natureza?"</i>
Output Típico	Alerta visual no monitor (<i>bounding box</i> , círculos ou sobreposições coloridas)	Classificação diagnóstica ou <i>score</i> de probabilidade
Exemplo em Modalidades Endoscópicas	ENDOBRRAIN-EYE™ Deteção automática de lesões em tempo real durante colonoscopia para maximizar a taxa de deteção de adenomas ³⁵	ENDOBRRAIN™ Recorre a imagens de endocitoscopia para auxiliar na distinção entre pólipos neoplásicos e não neoplásicos ¹⁵⁹

CAdE: *Computer-Aided Detection*; **CAdx:** *Computer-Aided Diagnosis*.

Tabela III — Tabela Comparativa Entre Biomarcadores Tradicionais e Digitais na DII

Baseada em ^{132,146}.

Característica	Biomarcadores Tradicionais	Biomarcadores Digitais
Colheita de Dados	Episódica (contexto clínico/hospitalar)	Contínua (ambiente real, monitorização remota)
Invasividade	Moderada a elevada (análises de sangue, fezes, endoscopia)	Não invasiva (<i>wearables</i> , apps, sensores remotos)
Sobrecarga do Doente	Elevada (deslocações e procedimentos)	Baixa (recolha passiva ou integrada na rotina)
Tempo de Resposta	Diferido (depende de consultas e tempos laboratoriais)	Imediato ou em tempo real
Capacidade Preditiva	Reativa (reflete atividade inflamatória atual)	Preditiva (alterações detetáveis até 7 semanas antes de <i>flare</i>)
Correlação Clínica	Bem estabelecida (padrão de referência)	Emergente , mas promissora (correlação com PCR e CPF)
Potencial de Personalização	Moderado	Elevado (ajuste dinâmico ao perfil do doente)
Exemplos	PCR, CPF, Lactoferrina	FC, VFC, SpO2 , marcadores no suor, padrões de sono

CPF: Calprotectina Fecal; FC: Frequência Cardíaca; VFC: Variabilidade da Frequência Cardíaca; SpO2: Saturação Periférica de O2; PCR: Proteína C Reativa.

INTELIGÊNCIA ARTIFICIAL (IA)

Sistemas capazes resolver problemas que tipicamente requerem inteligência humana.

MACHINE LEARNING (ML)

Algoritmos capazes de analisar dados para identificar padrões e completar tarefas sem serem especificamente programados para tal.

DEEP LEARNING (DL)

Modelos capazes de capturar relações mais intrincadas entre dados, possibilitando a abordagem de tarefas mais complexas.

Figura 1 — Esquematização dos Principais Subtipos de IA

Baseada em ^{7,8}.

(página intencionalmente deixada em branco)

FASE 1: DEFINIÇÃO DO PROBLEMA

PERGUNTA CLÍNICA E TAREFA COMPUTACIONAL

Discriminar mucosa normal vs. com lesão inflamatória → Classificação Supervisionada Binária

FASE 2: AQUISIÇÃO E CURADORIA DOS DADOS

RECOLHA DE IMAGENS

Frames de colonoscopias • Aprovação Ética
Anonimização dos Resultados

ANOTAÇÃO (*GROUND TRUTH*)

≥2 gastroenterologistas peritos • Arbitragem se discordância
Rótulo: lesão (1) vs. normal (0)

FASE 3: PRÉ-PROCESSAMENTO

NORMALIZAÇÃO, FILTRAGEM E *DATA AUGMENTATION*

Resolução, espaço de cor • exclusão de *frames* não-informativos • rotação, inversão, ajuste de brilho/contraste

FASE 4: DIVISÃO DO *DATASET*

TREINO (≈70%)

Ajuste de Pesos

VALIDAÇÃO (≈15%)

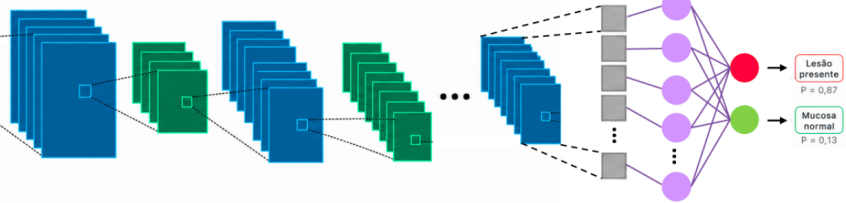
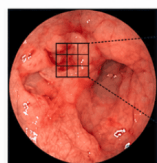
Hiperparâmetros • Early Stopping

TESTE INTERNO (≈15%)

Avaliação do Desempenho

FASE 5: DESENVOLVIMENTO E TREINO DO MODELO

CNN



Selecionar Arquitetura • Função de Perda • Hiperparâmetros

±Retreinamento

FASE 6: VALIDAÇÃO INTERNA

MÉTRICAS DE DISCRIMINAÇÃO

Sensibilidade, Especificidade, Exatidão, VPP, VPN, *F1-score*, AUC

FASE 7: VALIDAÇÃO EXTERNA E IMPLEMENTAÇÃO

VALIDAÇÃO EXTERNA

Datasets de centros independentes •
Avaliação da generalização • Ensaios clínicos

IMPLEMENTAÇÃO E MONITORIZAÇÃO

Integração em plataforma endoscópica (CAde) • vigilância
pós-implementação • re-treinamento do modelo

Figura 2 — Processo de Desenvolvimento e Avaliação de um Modelo de IA Hipotético

Esta figura corresponde a um cenário hipotético que ilustra os passos habitualmente tomados no desenvolvimento da maioria dos modelos mais comumente utilizados neste contexto.

*O primeiro passo (**fase 1**) corresponde à definição do problema e qual a tarefa computacional a aplicar. Neste caso, o objetivo seria discriminar entre mucosa normal e mucosa com lesão inflamatória, pelo que computacionalmente se pretende realizar uma classificação binária (normal [0] vs. lesão [1]) e será feita de modo supervisionado, ou seja, com inputs humanos no seu treino.*

*O próximo passo (**fase 2**) corresponde à aquisição dos dados e sua curadoria, um dos mais importantes, dado que não só é necessário um grande volume para um correto treinamento do modelo, como todo o seu treino e validação terá por base estes mesmos dados. Assim, neste caso particular, frames de colonoscopias afiguram-se excelentes candidatos a suprirem esta necessidade, pelo que foram escolhidos. Para o seu uso é necessário, obviamente, aprovação ética e anonimização dos resultados. Ainda nesta fase, os dados devem ser rotulados por humanos experientes para servir como ground truth pelos quais o modelo será treinado e relativamente aos quais os outputs do modelo serão avaliados. Esta rotulagem será feita para cada frame, indicando se observam lesão (1) ou não (2), numa avaliação dicotómica. Caso haja discordância, deve haver arbitragem para a resolver.*

*Após esta etapa, é necessário normalizar, filtrar e aumentar os dados (**fase 3**), de modo a standardizar os inputs do modelo. Algumas das operações mais utilizadas são padronizar a resolução das imagens e o seu aspeto geral, que pode beneficiar da exclusão de frames não informativos, ou de operações de rotação, inversão, ou alteração de brilho/contraste da imagem.*

*Uniformizado o dataset, este é dividido (**fase 4**), naquela que é, também, uma das etapas mais importantes do processo para garantir a validade do modelo. A divisão deve ser feita em três conjuntos: de treino ($\approx 70\%$), de validação ($\approx 15\%$) e teste interno ($\approx 15\%$). Os primeiros dois são utilizados no treino do modelo, enquanto o último é utilizado para realizar a validação interna do modelo, como descrito adiante. De notar que no conjunto de “validação”, este termo diz respeito ao conceito computacional e não estatístico, pelo que o conjunto utilizado na “validação interna” do modelo é efetivamente o de “teste interno” e não o de “validação”. Cada dataset corresponde, grosso modo, à correspondência entre imagem e lesão(1)/não(0).*

*Transversalmente a estas etapas, é necessário decidir qual a arquitetura de IA mais apropriada ao objetivo (**fase 5**). Embora recentemente novas arquiteturas mais complexas e capazes já existam, a mais utilizada para avaliação de imagens são as CNN (Convolution Neural Networks). Não obstante a descrição exata do seu funcionamento superar os propósitos desta dissertação, simplifadamente, esta arquitetura avalia porções de uma imagem às quais aplica filtros para retirar determinadas características. Estas são depois passadas para um conjunto de “neurónios” que pondera a importância de cada uma e como output apresenta a probabilidade de a lesão estar presente ou não. Para treinar estes modelos, recorre-se simultaneamente aos sets de treino (que permite aprimorar as ponderações referidas antes) e de validação (que permite otimizar parâmetros que, por sua vez, ditam como o algoritmo “deve aprender”).*

*Após esta aprendizagem, o modelo é validado internamente (**fase 6**) com base nas imagens do set de teste interno, que o modelo nunca utilizou no seu estudo, razão pela qual foi inicialmente “posto de parte” e só agora utilizado. Neste passo são avaliadas as várias métricas de desempenho como sensibilidade, especificidade, VPP, VPN, AUC, f1-score, que são depois publicados em artigos. Contudo, para avaliar a verdadeira aplicabilidade do modelo, é essencial efetuar a sua validação externa em dados multicêntricos (**fase 7**) através de ensaios clínicos robustos, que sustentem a sua posterior integração em plataformas a usar na prática médica. Nestas etapas, pode-se verificar que os resultados do modelo ficaram aquém do esperado, pelo que deverão ser retreinados (ou escolhida outra arquitetura, por exemplo).*

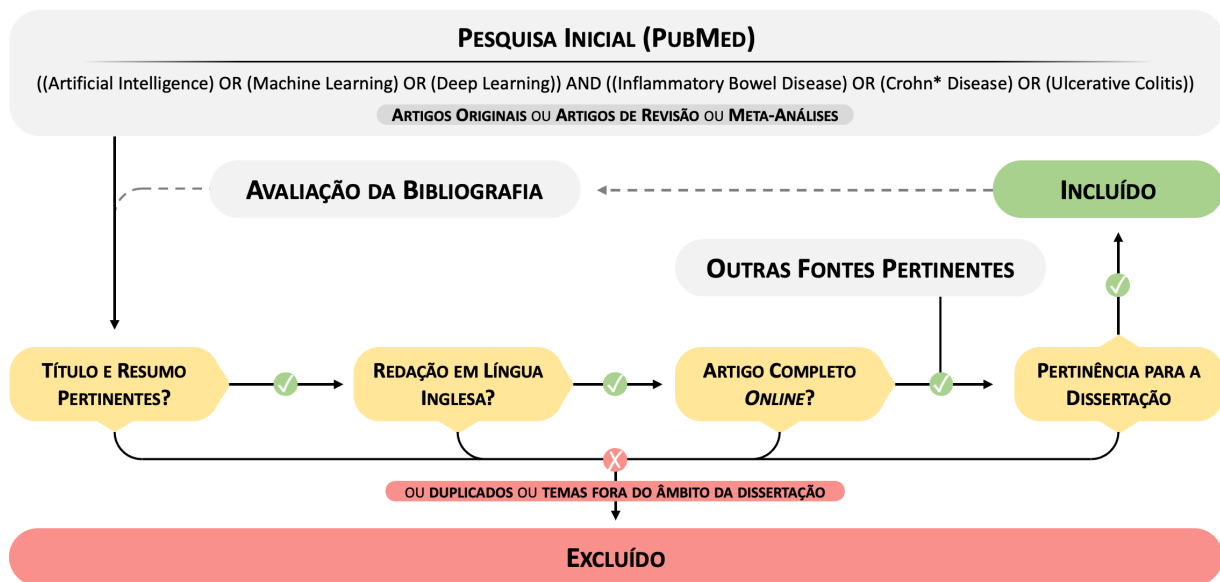


Figura 3 — Esquematização do Processo de Recolha e Análise Bibliográfica

BIBLIOGRAFIA

1. Dolinger M, Torres J, Vermeire S. Crohn's disease. *Lancet*. 2024 Mar 23;403(10432):1177–91.
2. Berre CL, Honap S, Peyrin-Biroulet L. Ulcerative colitis. *The Lancet*. 2023 Aug 12;402(10401):571–84.
3. Hashash JG, Limdi JK, Shapiro JM, Shah SA. Medical management of inflammatory bowel diseases. *BMJ*. 2025 Dec 19;391:e079050.
4. Kucharzik T, Taylor S, Allocca M, et al. ECCO-ESGAR-ESP-IBUS Guideline on Diagnostics and Monitoring of Patients with Inflammatory Bowel Disease: Part 1. *J Crohns Colitis*. 2025 July 3;19(7):jjaf106.
5. Kaplan GG. The global burden of inflammatory bowel disease: from 2025 to 2045. *Nat Rev Gastroenterol Hepatol*. 2025 July 18;
6. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med*. 2019 Jan;25(1):30–6.
7. Mota J, Almeida MJ, Martins M, et al. Artificial Intelligence in Coloproctology: A Review of Emerging Technologies and Clinical Applications. *J Clin Med*. 2024 Sept 30;13(19):5842.
8. Iacucci M, Santacroce G, Yasuharu M, Ghosh S. Artificial Intelligence–Driven Personalized Medicine: Transforming Clinical Practice in Inflammatory Bowel Disease. *Gastroenterology*. 2025 Aug 1;169(3):416–31.
9. Ruiqing L, Jing Y, Shunli L, et al. A Novel Radiomics Model Integrating Luminal and Mesenteric Features to Predict Mucosal Activity and Surgery Risk in Crohn's Disease Patients: A Multicenter Study. *Acad Radiol*. 2023 Sept;30 Suppl 1:S207–19.
10. Luseti F, Maimaris S, La Rosa GP, et al. Applications of generative artificial intelligence in inflammatory bowel disease: A systematic review. *Dig Liver Dis*. 2025 May;S1590865825007340.
11. Zhang Y, Wan X-H, Kong Q-Z, et al. Evaluating large language models as patient education tools for inflammatory bowel disease: A comparative study. *World J Gastroenterol*. 2025 Feb 14;31(6).
12. Kopylov U, Verstockt B, Marigorta UM, et al. Results of the 9th Scientific Workshop of the European Crohn's and Colitis Organisation (ECCO): Artificial Intelligence in medical management and precision medicine. *J Crohns Colitis*. 2025 Aug 12;jjaf134.
13. Banerjee R, Pal P, Mak JWY, Ng SC. Challenges in the diagnosis and management of inflammatory bowel disease in resource-limited settings in Asia. *Lancet Gastroenterol Hepatol*. 2020 Dec;5(12):1076–88.
14. Guimarães P, Finkler H, Reichert MC, et al. Artificial-intelligence-based decision support tools for the differential diagnosis of colitis. *Eur J Clin Invest*. 2023 June;53(6):e13960.
15. Luo X, Wang J, Tan C, et al. Rapid Endoscopic Diagnosis of Benign Ulcerative Colorectal Diseases With an Artificial Intelligence Contextual Framework. *Gastroenterology*. 2024 Aug 1;167(3):591–603.e9.
16. Yu Z, Liu X, Yu X, Xin Y, Zhou S, Li X. Development and application of an artificial intelligence-assisted endoscopic system for automatic and accurate diagnosis of colorectal ulcers. *Int J Colorectal Dis*. 2025 Dec 2;40(1):242.
17. Weng F, Meng Y, Lu F, et al. Differentiation of intestinal tuberculosis and Crohn's disease through an explainable machine learning method. *Sci Rep*. 2022 Feb 2;12(1):1714.

18. Gong T, Li M, Pu H, et al. Computed tomography enterography-based multiregional radiomics model for differential diagnosis of Crohn's disease from intestinal tuberculosis. *Abdom Radiol N Y*. 2023 June;48(6):1900–10.
19. Chang JT. Pathophysiology of Inflammatory Bowel Diseases. *N Engl J Med*. 2020 Dec 30;383(27):2652–64.
20. Wang X, Xiao Y, Xu X, et al. Characteristics of Fecal Microbiota and Machine Learning Strategy for Fecal Invasive Biomarkers in Pediatric Inflammatory Bowel Disease. *Front Cell Infect Microbiol*. 2021 Dec 7;11.
21. Ruan G, Qi J, Cheng Y, et al. Development and Validation of a Deep Neural Network for Accurate Identification of Endoscopic Images From Patients With Ulcerative Colitis and Crohn's Disease. *Front Med*. 2022;9:854677.
22. Quénéhervé L, David G, Bourreille A, et al. Quantitative assessment of mucosal architecture using computer-based analysis of confocal laser endomicroscopy in inflammatory bowel diseases. *Gastrointest Endosc*. 2019 Mar 1;89(3):626–36.
23. Mossotto E, Ashton JJ, Coelho T, Beattie RM, MacArthur BD, Ennis S. Classification of Paediatric Inflammatory Bowel Disease using Machine Learning. *Sci Rep*. 2017 May 25;7(1):2427.
24. Rymarczyk D, Schultz W, Borowa A, et al. Deep Learning Models Capture Histological Disease Activity in Crohn's Disease and Ulcerative Colitis with High Fidelity. *J Crohns Colitis*. 2024 Apr 23;18(4):604–14.
25. Furlanello C, Bussola N, Merzi N, et al. The development of artificial intelligence in the histological diagnosis of Inflammatory Bowel Disease (IBD-AI). *Dig Liver Dis Off J Ital Soc Gastroenterol Ital Assoc Study Liver*. 2025 Jan;57(1):184–9.
26. Cai C, Shi Q, Liu M, et al. IBDAIM:Artificial intelligence for analyzing intestinal biopsies pathological images for assisted integrated diagnostic of inflammatory bowel disease. *Int J Med Inf*. 2025 Nov 1;203:106024.
27. Zhou Z, Xiong Z, Cheng R, et al. Volumetric visceral fat machine learning phenotype on CT for differential diagnosis of inflammatory bowel disease. *Eur Radiol*. 2023 Mar;33(3):1862–72.
28. Pei J, Wang G, Li Y, et al. Utility of four machine learning approaches for identifying ulcerative colitis and Crohn's disease. *Heliyon*. 2024 Jan 15;10(1):e23439.
29. Wei Z, Wang W, Bradfield J, et al. Large Sample Size, Wide Variant Spectrum, and Advanced Machine-Learning Technique Boost Risk Prediction for Inflammatory Bowel Disease. *Am J Hum Genet*. 2013 June 6;92(6):1008–12.
30. Stafford IS, Ashton JJ, Mossotto E, Cheng G, Mark Beattie R, Ennis S. Supervised Machine Learning Classifies Inflammatory Bowel Disease Patients by Subtype Using Whole Exome Sequencing Data. *J Crohns Colitis*. 2023 Oct 1;17(10):1672–80.
31. Kim H, Na JE, Kim S, et al. A Machine Learning-Based Diagnostic Model for Crohn's Disease and Ulcerative Colitis Utilizing Fecal Microbiome Analysis. *Microorganisms*. 2023 Dec 24;12(1):36.

32. Tong Y, Lu K, Yang Y, et al. Can natural language processing help differentiate inflammatory intestinal diseases in China? Models applying random forest and convolutional neural network approaches. *BMC Med Inform Decis Mak*. 2020 Sept 29;20(1):248.
33. Turner D, Ricciuto A, Lewis A, et al. STRIDE-II: An Update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) Initiative of the International Organization for the Study of IBD (IOIBD): Determining Therapeutic Goals for Treat-to-Target strategies in IBD. *Gastroenterology*. 2021 Apr;160(5):1570–83.
34. Hassan C, Bisschops R, Sharma P, Mori Y. Colon Cancer Screening, Surveillance, and Treatment: Novel Artificial Intelligence Driving Strategies in the Management of Colon Lesions. *Gastroenterology*. 2025 Aug 1;169(3):444–55.
35. Misawa M, Kudo S, Mori Y, et al. Development of a computer-aided detection system for colonoscopy and a publicly accessible large colonoscopy video database (with video). *Gastrointest Endosc*. 2021 Apr 1;93(4):960-967.e3.
36. Stidham RW, Liu W, Bishu S, et al. Performance of a Deep Learning Model vs Human Reviewers in Grading Endoscopic Disease Severity of Patients With Ulcerative Colitis. *JAMA Netw Open*. 2019 May 17;2(5):e193963.
37. Takenaka K, Ohtsuka K, Fujii T, et al. Development and Validation of a Deep Neural Network for Accurate Evaluation of Endoscopic Images From Patients With Ulcerative Colitis. *Gastroenterology*. 2020 June 1;158(8):2150–7.
38. Gottlieb K, Requa J, Karnes W, et al. Central Reading of Ulcerative Colitis Clinical Trial Videos Using Neural Networks. *Gastroenterology*. 2021 Feb 1;160(3):710-719.e2.
39. Fan Y, Mu R, Xu H, et al. Novel deep learning-based computer-aided diagnosis system for predicting inflammatory activity in ulcerative colitis. *Gastrointest Endosc*. 2023 Feb;97(2):335–46.
40. Yao H, Najarian K, Gryak J, et al. Fully automated endoscopic disease activity assessment in ulcerative colitis. *Gastrointest Endosc*. 2021 Mar 1;93(3):728-736.e1.
41. Qi J, Ruan G, Ping Y, et al. Development and validation of a deep learning-based approach to predict the Mayo endoscopic score of ulcerative colitis. *Ther Adv Gastroenterol*. 2023 May 22;16:17562848231170945.
42. Byrne MF, Panaccione R, East JE, et al. Application of Deep Learning Models to Improve Ulcerative Colitis Endoscopic Disease Activity Scoring Under Multiple Scoring Systems. *J Crohns Colitis*. 2023 Apr 1;17(4):463–71.
43. Jahagirdar V, Bapaye J, Chandan S, et al. Diagnostic accuracy of convolutional neural network-based machine learning algorithms in endoscopic severity prediction of ulcerative colitis: a systematic review and meta-analysis. *Gastrointest Endosc*. 2023 Aug;98(2):145-154.e8.
44. Stidham RW, Cai L, Cheng S, et al. Using Computer Vision to Improve Endoscopic Disease Quantification in Therapeutic Clinical Trials of Ulcerative Colitis. *Gastroenterology*. 2024 Jan;166(1):155-167.e2.

45. Gutierrez-Becker B, Fraessle S, Yao H, et al. Ulcerative Colitis Severity Classification and Localized Extent (UC-SCALE): An Artificial Intelligence Scoring System for a Spatial Assessment of Disease Severity in Ulcerative Colitis. *J Crohns Colitis*. 2025 Jan 1;19(1):jjae187.
46. Mascarenhas M, Afonso J, Andrade P, Cardoso H, Macedo G. Artificial intelligence and capsule endoscopy: unravelling the future. *Ann Gastroenterol*. 2021;34(3):300–9.
47. Aoki T, Yamada A, Aoyama K, et al. Clinical usefulness of a deep learning-based system as the first screening on small-bowel capsule endoscopy reading. *Dig Endosc*. 2020;32(4):585–91.
48. Oh DJ, Hwang Y, Kim SH, Nam JH, Jung MK, Lim YJ. Reading of small bowel capsule endoscopy after frame reduction using an artificial intelligence algorithm. *BMC Gastroenterol*. 2024 Feb 22;24(1):80.
49. Brodersen JB, Jensen MD, Leenhardt R, et al. Artificial Intelligence-assisted Analysis of Pan-enteric Capsule Endoscopy in Patients with Suspected Crohn’s Disease: A Study on Diagnostic Performance. *J Crohns Colitis*. 2024 Jan 27;18(1):75–81.
50. Klang E, Barash Y, Margalit RY, et al. Deep learning algorithms for automated detection of Crohn’s disease ulcers by video capsule endoscopy. *Gastrointest Endosc*. 2020 Mar;91(3):606-613.e2.
51. Klang E, Grinman A, Soffer S, et al. Automated Detection of Crohn’s Disease Intestinal Strictures on Capsule Endoscopy Images Using Deep Neural Networks. *J Crohns Colitis*. 2021 May 4;15(5):749–56.
52. Klang E, Kopylov U, Mortensen B, et al. A Convolutional Neural Network Deep Learning Model Trained on CD Ulcers Images Accurately Identifies NSAID Ulcers. *Front Med*. 2021;8:656493.
53. Barash Y, Azaria L, Soffer S, et al. Ulcer severity grading in video capsule images of patients with Crohn’s disease: an ordinal neural network solution. *Gastrointest Endosc*. 2021 Jan 1;93(1):187–92.
54. Giordano A, Romero-Mascarell C, González-Suárez B, Guarner-Argente C. Integration of Artificial Intelligence-Enhanced Capsule Endoscopy in Clinical Practice: A Review of Market-Available Tools for Clinical Practice. *Dig Dis Sci*. 2025 Sept;70(9):2966–76.
55. Cardoso P, Mascarenhas M, Afonso J, et al. Deep learning and minimally invasive inflammatory activity assessment: a proof-of-concept study for development and score correlation of a panendoscopy convolutional network. *Ther Adv Gastroenterol*. 2024;17:17562848241251569.
56. Freitas M, Arieira C, Carvalho PB, Rosa B, Moreira MJ, Cotter J. Simplify to improve in capsule endoscopy - TOP 100 is a swift and reliable evaluation tool for the small bowel inflammatory activity in Crohn’s disease. *Scand J Gastroenterol*. 2020 Apr;55(4):408–13.
57. Ferreira JPS, de Mascarenhas Saraiva MJ da QEC, Afonso JPL, et al. Identification of Ulcers and Erosions by the Novel Pillcam™ Crohn’s Capsule Using a Convolutional Neural Network: A Multicentre Pilot Study. *J Crohns Colitis*. 2022 Jan 28;16(1):169–72.
58. Andrade P, Mascarenhas M, Mendes F, et al. AI-Assisted Capsule Endoscopy for Detection of Ulcers and Erosions in Crohn’s Disease: A Multicenter Validation Study. *Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc*. 2025 Nov 5;S1542-3565(25)00861-4.
59. Mesquita P, Pinho R, Rodrigues A, et al. OMOM SmartScan provides noninferior assessment of Lewis score and CECDAI in Crohn’s disease vs full-video review. *Dig Liver Dis Off J Ital Soc Gastroenterol Ital Assoc Study Liver*. 2026 Mar;58(3):377–83.

60. Fan D, Mao Y, Liang F, et al. INTELCAPE: A Deep Learning-Powered System for Automated, High-Accuracy Crohn's Disease Diagnosis via Capsule Endoscopy. *Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc.* 2026 Mar 19;S1542-3565(26)00184-9.
61. Jiang X, Luo X, Nan Q, Ye Y, Miao Y, Miao J. Application of deep learning in the diagnosis and evaluation of ulcerative colitis disease severity. *Ther Adv Gastroenterol.* 2023;16:17562848231215579.
62. Bossuyt P, Nakase H, Vermeire S, et al. Automatic, computer-aided determination of endoscopic and histological inflammation in patients with mild to moderate ulcerative colitis based on red density. *Gut.* 2020 Oct 1;69(10):1778–86.
63. Bossuyt P, De Hertogh G, Eelbode T, Vermeire S, Bisschops R. Computer-Aided Diagnosis With Monochromatic Light Endoscopy for Scoring Histologic Remission in Ulcerative Colitis. *Gastroenterology.* 2021 Jan;160(1):23–5.
64. Sinonquel P, Lenfant M, Eelbode T, et al. Development of an Automated Tool for the Estimation of Histological Remission in Ulcerative Colitis Using Single-Wavelength Endoscopy Technology. *J Crohns Colitis.* 2025 May 1;19(5):jjae180.
65. Maeda Y, Kudo S, Mori Y, et al. Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video). *Gastrointest Endosc.* 2019 Feb 1;89(2):408–15.
66. Iacucci M, Jeffery L, Acharjee A, et al. Ultra-high Magnification Endocytoscopy and Molecular Markers for Defining Endoscopic and Histologic Remission in Ulcerative Colitis-An Exploratory Study to Define Deep Remission. *Inflamm Bowel Dis.* 2021 Oct 20;27(11):1719–30.
67. Udristoiu AL, Stefanescu D, Gruionu G, et al. Deep Learning Algorithm for the Confirmation of Mucosal Healing in Crohn's Disease, Based on Confocal Laser Endomicroscopy Images. *J Gastrointest Liver Dis JGLD.* 2021 Mar 12;30(1):59–65.
68. Vande Casteele N, Leighton JA, Pasha SF, et al. Utilizing Deep Learning to Analyze Whole Slide Images of Colonic Biopsies for Associations Between Eosinophil Density and Clinicopathologic Features in Active Ulcerative Colitis. *Inflamm Bowel Dis.* 2022 Apr 1;28(4):539–46.
69. Gui X, Bazarova A, Del Amor R, et al. PICaSSO Histologic Remission Index (PHRI) in ulcerative colitis: development of a novel simplified histological score for monitoring mucosal healing and predicting clinical outcomes and its applicability in an artificial intelligence system. *Gut.* 2022 May;71(5):889–98.
70. Najdawi F, Sucipto K, Mistry P, et al. Artificial Intelligence Enables Quantitative Assessment of Ulcerative Colitis Histology. *Mod Pathol Off J U S Can Acad Pathol Inc.* 2023 June;36(6):100124.
71. Peyrin-Biroulet L, Adsul S, Stancati A, Dehmeshki J, Kubassova O. An artificial intelligence-driven scoring system to measure histological disease activity in ulcerative colitis. *United Eur Gastroenterol J.* 2024 Oct;12(8):1028–33.
72. Mookhoek A, Sinonque P, Allocca M, et al. Results of the 9th Scientific Workshop of the European Crohn's and Colitis Organisation (ECCO): Artificial Intelligence in Endoscopy, Radiology and Histology in IBD Diagnostics. *J Crohns Colitis.* 2025 Aug 12;jjaf133.

73. Carter D, Albshesh A, Shimon C, et al. Automatized Detection of Crohn's Disease in Intestinal Ultrasound Using Convolutional Neural Network. *Inflamm Bowel Dis*. 2023 Dec 5;29(12):1901–6.
74. Kumaralingam L, Le May K, Dang VB, Alavi J, Huynh HQ, Le LH. Artificial intelligence-assisted approach to assessing bowel wall thickness in pediatric inflammatory bowel disease using intestinal ultrasound images. *J Crohns Colitis*. 2025 Apr 1;19(4):jjaf037.
75. Gu P, Chang J-H, Carter D, et al. Radiomics-Based Analysis of Intestinal Ultrasound Images for Inflammatory Bowel Disease: A Feasibility Study. *Crohns Colitis* 360. 2024 Apr;6(2):otae034.
76. Naziroglu RE, Puylaert CAJ, Tielbeek JAW, et al. Semi-automatic bowel wall thickness measurements on MR enterography in patients with Crohn's disease. *Br J Radiol*. 2017 June;90(1074):20160654.
77. Kneib A, Lemoire A, Hossu G, Peyrin-Biroulet L, Laurent V, Odille F. Semi-Automatic Assessment of Crohn's Disease Activity by Combined Analysis of Bowel Lesions and Creeping Fat. *Magn Reson Med*. 2026 Mar 7;
78. Chirra P, Sleiman J, Gandhi NS, et al. Radiomics to Detect Inflammation and Fibrosis on Magnetic Resonance Enterography in Stricturing Crohn's Disease. *J Crohns Colitis*. 2024 Oct 15;18(10):1660–71.
79. Arkko A, Kaseva T, Salli E, Mäkelä T, Savolainen S, Kangasniemi M. Automatic detection of Crohn's disease using quantified motility in magnetic resonance enterography: initial experiences. *Clin Radiol*. 2022 Feb;77(2):96–103.
80. Gollifer RM, Menys A, Plumb A, et al. Automated versus subjective assessment of spatial and temporal MRI small bowel motility in Crohn's disease. *Clin Radiol*. 2019 Oct;74(10):814.e9-814.e19.
81. Ding H, Li J, Jiang K, et al. Assessing the inflammatory severity of the terminal ileum in Crohn disease using radiomics based on MRI. *BMC Med Imaging*. 2022 July 4;22(1):118.
82. Guez I, Focht G, Greer M-LC, et al. Development of a multimodal machine-learning fusion model to non-invasively assess ileal Crohn's disease endoscopic activity. *Comput Methods Programs Biomed*. 2022 Dec;227:107207.
83. Yuan G, He Y, Cao Q-H, et al. Visceral adipose volume is correlated with surgical tissue fibrosis in Crohn's disease of the small bowel. *Gastroenterol Rep*. 2022;10:goac044.
84. Li X, Liang D, Meng J, et al. Development and Validation of a Novel Computed-Tomography Enterography Radiomic Approach for Characterization of Intestinal Fibrosis in Crohn's Disease. *Gastroenterology*. 2021 June;160(7):2303-2316.e11.
85. Meng J, Luo Z, Chen Z, et al. Intestinal fibrosis classification in patients with Crohn's disease using CT enterography-based deep learning: comparisons with radiomics and radiologists. *Eur Radiol*. 2022 Dec;32(12):8692–705.
86. Li T, Liu Y, Guo J, Wang Y. Prediction of the activity of Crohn's disease based on CT radiomics combined with machine learning models. *J X-Ray Sci Technol*. 2022;30(6):1155–68.
87. Iacucci M, Parigi TL, Amor RD, et al. Artificial Intelligence Enabled Histological Prediction of Remission or Activity and Clinical Outcomes in Ulcerative Colitis. *Gastroenterology*. 2023 June 1;164(7):1180-1188.e2.

88. Ohara J, Nemoto T, Maeda Y, Ogata N, Kudo S-E, Yamochi T. Deep learning-based automated quantification of goblet cell mucus using histological images as a predictor of clinical relapse of ulcerative colitis with endoscopic remission. *J Gastroenterol*. 2022 Dec;57(12):962–70.
89. Iacucci M, Majumder S, Zammarchi I, et al. Automated real-time imaging of intestinal barrier integrity and molecular profiling for early outcome prediction in inflammatory bowel disease: endo-histo-barrier-omics study. *J Crohns Colitis*. 2025 Nov 1;19(11):jjaf200.
90. Fiorino G, Danese S, Peyrin-Biroulet L, et al. LIAISON® Calprotectin for the prediction of relapse in quiescent ulcerative colitis: The EuReCa study. *United Eur Gastroenterol J*. 2022 Oct;10(8):836–43.
91. Gan RW, Sun D, Tatro AR, et al. Replicating prediction algorithms for hospitalization and corticosteroid use in patients with inflammatory bowel disease. *PloS One*. 2021;16(9):e0257520.
92. Siegel CA, Siegel LS, Dubinsky MC, et al. Performance Characteristics of a Clinical Decision Support Tool for Disease Complications in Crohn’s Disease. *Crohns Colitis* 360. 2021 Oct;3(4):otab074.
93. Ungaro RC, Hu L, Ji J, et al. Machine learning identifies novel blood protein predictors of penetrating and stricturing complications in newly diagnosed paediatric Crohn’s disease. *Aliment Pharmacol Ther*. 2021;53(2):281–90.
94. Kurowski JA, Achkar J-P, Sugano D, et al. Computable Phenotype of a Crohn’s Disease Natural History Model. *Med Decis Mak Int J Soc Med Decis Mak*. 2022 Oct;42(7):937–44.
95. Dalal RS, Osterman MT, Buchner AM, Praestgaard A, Lewis JD, Lichtenstein GR. A User-Friendly Prediction Tool to Identify Colectomy Risk in Patients With Ulcerative Colitis. *Inflamm Bowel Dis*. 2019 Aug 20;25(9):1550–8.
96. Baut GL, Kirchgessner J, Amiot A, et al. A Scoring System to Determine Patients’ Risk of Colectomy Within 1 Year After Hospital Admission for Acute Severe Ulcerative Colitis. *Clin Gastroenterol Hepatol*. 2021 Aug 1;19(8):1602-1610.e1.
97. Magro F, Dias CC, Portela F, et al. Development and Validation of Risk Matrices Concerning Ulcerative Colitis Outcomes-Bayesian Network Analysis. *J Crohns Colitis*. 2019 Mar 30;13(4):401–9.
98. Stidham RW, Liu Y, Enchalody B, et al. The Use of Readily Available Longitudinal Data to Predict the Likelihood of Surgery in Crohn Disease. *Inflamm Bowel Dis*. 2021 Aug 1;27(8):1328–34.
99. Chirra P, Sharma A, Bera K, et al. Integrating Radiomics With Clinicoradiological Scoring Can Predict High-Risk Patients Who Need Surgery in Crohn’s Disease: A Pilot Study. *Inflamm Bowel Dis*. 2023 Mar 1;29(3):349–58.
100. Ochs V, Taha-Mehlitz S, Lavanchy JL, et al. Development of a Clinical Prediction Model for Anastomotic Leakage in Colorectal Surgery. *JAMA Netw Open*. 2025 Oct 20;8(10):e2538267.
101. Mascarenhas M, Mendes F, Fonseca F, et al. A Novel Deep Learning Model for Predicting Colorectal Anastomotic Leakage: A Pioneer Multicenter Transatlantic Study. *J Clin Med*. 2025 Aug 3;14(15):5462.
102. Allez M, Bak MTJ, Brand S, et al. Development and external validation of a predictive model for postoperative recurrence of Crohn’s disease in the biologic era. *Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc*. 2026 Mar 27;S1542-3565(26)00220-X.

- 103.Kiyokawa H, Abe M, Matsui T, et al. Deep Learning Analysis of Histologic Images from Intestinal Specimen Reveals Adipocyte Shrinkage and Mast Cell Infiltration to Predict Postoperative Crohn Disease. *Am J Pathol*. 2022 June;192(6):904–16.
- 104.Wang Y, He Q, Yao D, et al. Histological Image-based Ensemble Model to Identify Myenteric Plexitis and Predict Endoscopic Postoperative Recurrence in Crohn's Disease: A Multicentre, Retrospective Study. *J Crohns Colitis*. 2024 May 31;18(5):727–37.
- 105.Moran GW, Gordon M, Sinopoulou V, et al. British Society of Gastroenterology guidelines on inflammatory bowel disease in adults: 2025. *Gut*. 2025 June 1;74(Suppl 2):s1–101.
- 106.Schöler D, Kostev K, Peters M, et al. Machine Learning Can Predict the Probability of Biologic Therapy in Patients with Inflammatory Bowel Disease. *J Clin Med*. 2022 Aug 5;11(15):4586.
- 107.Kellerman R, Bleiweiss A, Samuel S, et al. Spatiotemporal analysis of small bowel capsule endoscopy videos for outcomes prediction in Crohn's disease. *Ther Adv Gastroenterol*. 2023;16:17562848231172556.
- 108.Alsoud D, Sabino J, Ferrante M, Verstockt B, Vermeire S. Calibration, Clinical Utility, and Specificity of Clinical Decision Support Tools in Inflammatory Bowel Disease. *Clin Gastroenterol Hepatol*. 2025 June 1;23(7):1216-1227.e14.
- 109.Dulai PS, Amiot A, Peyrin-Biroulet L, et al. A clinical decision support tool may help to optimise vedolizumab therapy in Crohn's disease. *Aliment Pharmacol Ther*. 2020;51(5):553–64.
- 110.Shi Y, He W, Zhong M, Yu M. MIN score predicts primary response to infliximab/adalimumab and vedolizumab therapy in patients with inflammatory bowel diseases. *Genomics*. 2021 July;113(4):1988–98.
- 111.Waljee AK, Sauder K, Patel A, et al. Machine Learning Algorithms for Objective Remission and Clinical Outcomes with Thiopurines. *J Crohns Colitis*. 2017 July 1;11(7):801–10.
- 112.Frick S, Müller D, Kullak-Ublick GA, Jetter A. Do clinical and laboratory parameters predict thiopurine metabolism and clinical outcome in patients with inflammatory bowel diseases? *Eur J Clin Pharmacol*. 2019 Mar;75(3):335–42.
- 113.Liu C, Zhang Y, Wang Y-J, et al. The potential of lymphocyte-C-reactive protein ratio upon admission to predict therapeutic responsiveness to mesalazine in adult patients with mild-to-moderate active ulcerative colitis. *BMC Gastroenterol*. 2025 Nov 28;25(1):895.
- 114.Ma C, Jeyarajah J, Guizzetti L, et al. Modeling Endoscopic Improvement after Induction Treatment With Mesalamine in Patients With Mild-to-Moderate Ulcerative Colitis. *Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc*. 2022 Feb;20(2):447-454.e1.
- 115.Con D, Langenberg DR van, Vasudevan A. Deep learning vs conventional learning algorithms for clinical prediction in Crohn's disease: A proof-of-concept study. *World J Gastroenterol*. 2021 Oct 14;27(38):6476–88.
- 116.Li Y, Pan J, Zhou N, et al. A random forest model predicts responses to infliximab in Crohn's disease based on clinical and serological parameters. *Scand J Gastroenterol*. 2021 Sept 2;56(9):1030–9.

117. Qiu Y, Hu S, Chao K, et al. Developing a Machine-Learning Prediction Model for Infliximab Response in Crohn's Disease: Integrating Clinical Characteristics and Longitudinal Laboratory Trends. *Inflamm Bowel Dis*. 2025 May 1;31(5):1334–43.
118. Feng J, Feng Q, Chen Y, et al. MRI-Based Radiomic Signature Identifying Secondary Loss of Response to Infliximab in Crohn's Disease. *Front Nutr*. 2021;8:773040.
119. Zhu C, Hu J, Wang X, et al. A novel clinical radiomics nomogram at baseline to predict mucosal healing in Crohn's disease patients treated with infliximab. *Eur Radiol*. 2022 Oct 1;32(10):6628–36.
120. Park SK, Kim YB, Kim S, et al. Development of a Machine Learning Model to Predict Non-Durable Response to Anti-TNF Therapy in Crohn's Disease Using Transcriptome Imputed from Genotypes. *J Pers Med*. 2022 June;12(6):947.
121. Mishra N, Aden K, Blase JI, et al. Longitudinal multi-omics analysis identifies early blood-based predictors of anti-TNF therapy response in inflammatory bowel disease. *Genome Med*. 2022 Sept 24;14(1):110.
122. Miyoshi J, Maeda T, Matsuoka K, et al. Machine learning using clinical data at baseline predicts the efficacy of vedolizumab at week 22 in patients with ulcerative colitis. *Sci Rep*. 2021 Aug 12;11(1):16440.
123. Pugliese D, Privitera G, Cersullo N, et al. Vedolizumab in inflammatory bowel disease: Real-world outcomes and their prediction with machine learning—the IG-IBD LIVE study. *Dig Liver Dis*. 2025 July 1;57(7):1393–402.
124. Venkatapurapu SP, Iwakiri R, Udagawa E, et al. A Computational Platform Integrating a Mechanistic Model of Crohn's Disease for Predicting Temporal Progression of Mucosal Damage and Healing. *Adv Ther*. 2022 July 1;39(7):3225–47.
125. Abreu MT, Davies JM, Quintero MA, et al. Transcriptional Behavior of Regulatory T Cells Predicts IBD Patient Responses to Vedolizumab Therapy. *Inflamm Bowel Dis*. 2022 Dec 1;28(12):1800–12.
126. Caenepeel C, Falony G, Machiels K, et al. Dysbiosis and Associated Stool Features Improve Prediction of Response to Biological Therapy in Inflammatory Bowel Disease. *Gastroenterology*. 2024 Mar 1;166(3):483–95.
127. Waljee AK, Wallace BI, Cohen-Mekelburg S, et al. Development and Validation of Machine Learning Models in Prediction of Remission in Patients With Moderate to Severe Crohn Disease. *JAMA Netw Open*. 2019 May 10;2(5):e193721.
128. Guo M, Guan Y, Hu S, et al. CT-based radiomics of bowel wall at baseline predicts the efficacy of Ustekinumab at week 16 in patients with Crohn's disease. *Abdom Radiol*. 2026 June;51(6):2764–75.
129. He M, Li C, Tang W, Kang Y, Zuo Y, Wang Y. Machine learning gene expression predicting model for ustekinumab response in patients with Crohn's disease. *Immun Inflamm Dis*. 2021;9(4):1529–40.
130. Harun R, Lu J, Kassir N, Zhang W. Machine Learning-Based Quantification of Patient Factors Impacting Remission in Patients With Ulcerative Colitis: Insights from Etrolizumab Phase III Clinical Trials. *Clin Pharmacol Ther*. 2024;115(4):815–24.

131. Iacucci M, Santacroce G, Meseguer P, et al. Endo-Histo foundational fusion model: a novel artificial intelligence for assessing histologic remission and response to therapy in ulcerative colitis clinical trial. *J Crohns Colitis*. 2025 July 1;19(7):jjaf108.
132. De Deo D, Dal Buono A, Gabbiadini R, et al. Digital biomarkers and artificial intelligence: a new frontier in personalized management of inflammatory bowel disease. *Front Immunol*. 2025 Aug 4;16:1637159.
133. Shinozaki S, Watanabe J, Kanno T, Nakazawa K, Yano T. Comparative Performance of Artificial Intelligence-Based Computer-Aided Detection Systems for Colorectal Polyps: A Systematic Review and Network Meta-Analysis. *Dig Endosc*. 2026 Apr;38(4):e70138.
134. Bretthauer M, Ahmed J, Antonelli G, et al. Use of computer-assisted detection (CAdE) colonoscopy in colorectal cancer screening and surveillance: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. *Endoscopy*. 2025 June;57(6):667–73.
135. Sultan S, Shung DL, Kolb JM, et al. AGA Living Clinical Practice Guideline on Computer-Aided Detection–Assisted Colonoscopy. *Gastroenterology*. 2025 Apr 1;168(4):691–700.
136. López-Serrano A, Voces A, Lorente JR, et al. Artificial intelligence for dysplasia detection during surveillance colonoscopy in patients with ulcerative colitis: A cross-sectional, non-inferiority, diagnostic test comparison study. *Gastroenterol Hepatol*. 2025 Feb;48(2):502210.
137. Fukunaga S, Kusaba Y, Ohuchi A, et al. Is artificial intelligence a superior diagnostician in ulcerative colitis? *Endoscopy*. 2021 Feb;53(02):E75–6.
138. Maeda Y, Kudo S-E, Ogata N, et al. Can artificial intelligence help to detect dysplasia in patients with ulcerative colitis? *Endoscopy*. 2021 July;53(7):E273–4.
139. Vinsard DG, Fetzer JR, Agrawal U, et al. Development of an artificial intelligence tool for detecting colorectal lesions in inflammatory bowel disease. *iGIE*. 2023 June 1;2(2):91-101.e6.
140. Yamamoto S, Kinugasa H, Hamada K, et al. The diagnostic ability to classify neoplasias occurring in inflammatory bowel disease by artificial intelligence and endoscopists: A pilot study. *J Gastroenterol Hepatol*. 2022;37(8):1610–6.
141. Noguchi T, Ando T, Emoto S, et al. Artificial Intelligence Program to Predict p53 Mutations in Ulcerative Colitis-Associated Cancer or Dysplasia. *Inflamm Bowel Dis*. 2022 July 1;28(7):1072–80.
142. Carreras J, Roncador G, Hamoudi R. Ulcerative Colitis, LAIR1 and TOX2 Expression, and Colorectal Cancer Deep Learning Image Classification Using Convolutional Neural Networks. *Cancers*. 2024 Dec 19;16(24):4230.
143. Burton AM, Perler BK. Mobile health technologies in inflammatory bowel disease: a narrative review. *BMC Gastroenterol*. 2025 Aug 18;25:595.
144. Del Hoyo J, Nos P, Faubel R, et al. A Web-Based Telemanagement System for Improving Disease Activity and Quality of Life in Patients With Complex Inflammatory Bowel Disease: Pilot Randomized Controlled Trial. *J Med Internet Res*. 2018 Nov 27;20(11):e11602.
145. Okegunna Y, Sekar M, Adriaans GMC, et al. Advancing Predictive Modeling of Inflammatory Bowel Disease (IBD) Flares: A Data-Driven Approach Using Lifestyle and Psychosocial Factors from a Remote Monitoring Platform. *Dig Dis Sci*. 2026 Mar 12;

146. Hirten RP, Danieleto M, Sanchez-Mayor M, et al. Physiological Data Collected From Wearable Devices Identify and Predict Inflammatory Bowel Disease Flares. *Gastroenterology*. 2025 May 1;168(5):939-951.e5.
147. Jagannath B, Lin K-C, Pali M, Sankhala D, Muthukumar S, Prasad S. A Sweat-based Wearable Enabling Technology for Real-time Monitoring of IL-1 β and CRP as Potential Markers for Inflammatory Bowel Disease. *Inflamm Bowel Dis*. 2020 Sept 18;26(10):1533–42.
148. Shahub S, Kumar RM, Lin K-C, et al. Continuous Monitoring of CRP, IL-6, and Calprotectin in Inflammatory Bowel Disease Using a Perspiration-Based Wearable Device. *Inflamm Bowel Dis*. 2025 Mar 1;31(3):647–54.
149. Honap S, Jairath V, Danese S, Peyrin-Biroulet L. Navigating the complexities of drug development for inflammatory bowel disease. *Nat Rev Drug Discov*. 2024 July;23(7):546–62.
150. Lahat A, Shachar E, Avidan B, Shatz Z, Glicksberg BS, Klang E. Evaluating the use of large language model in identifying top research questions in gastroenterology. *Sci Rep*. 2023 Mar 13;13(1):4164.
151. Stammers M, Sartain S, Cummings JRF, et al. Identification of Cohorts with Inflammatory Bowel Disease Amidst Fragmented Clinical Databases via Machine Learning. *Dig Dis Sci*. 2025 Oct;70(10):3309–22.
152. Stammers M, Gwiggner M, Nouraei R, Metcalf C, Batchelor J. Robust comparative evaluation of 15 natural language processing algorithms to positively identify patients with inflammatory bowel disease from secondary care records. *BMJ Open Gastroenterol*. 2025 Oct 10;12(1):e001977.
153. Jiang X, Lu L, Li J, et al. Synthetically Feasible De Novo Molecular Design of Leads Based on a Reinforcement Learning Model: AI-Assisted Discovery of an Anti-IBD Lead Targeting CXCR4. *J Med Chem*. 2024 June 27;67(12):10057–75.
154. Đogatović K, Vučićević K, Kovačević M, Ćulafić M, Miljković B, Kovačević SV. Identification of Drug Repurposing Opportunities of Immunomodulatory Drugs for Inflammatory Bowel Disease Through Inverse Pharmacovigilance Signal Detection in the FAERS Database. *J Clin Med*. 2026 Mar 12;15(6):2172.
155. Liu X, Reigle J, Prasath VBS, Dhaliwal J. Artificial intelligence image-based prediction models in IBD exhibit high risk of bias: A systematic review. *Comput Biol Med*. 2024 Mar 1;171:108093.
156. Noor NM, Daperno M, de Laffolie J, et al. Results of the 9th Scientific Workshop of the European Crohn's and Colitis Organisation (ECCO): Artificial Intelligence in IBD: Regulatory and Methodological Considerations. *J Crohns Colitis*. 2025 Aug 12;jjaf136.
157. Leggett CL, Parasa S, Repici A, Berzin TM, Gross SA, Sharma P. Physician perceptions on the current and future impact of artificial intelligence to the field of gastroenterology. *Gastrointest Endosc*. 2024 Apr 1;99(4):483-489.e2.
158. Sairenji T, Collins KL, Evans DV. An Update on Inflammatory Bowel Disease. *Prim Care Clin Off Pract*. 2017 Dec 1;44(4):673–92.
159. Barua I, Wieszczy P, Kudo S, et al. Real-Time Artificial Intelligence–Based Optical Diagnosis of Neoplastic Polyps during Colonoscopy. *NEJM Evid*. 2022 May 24;1(6):EVIDoa2200003.

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR

