

EDITORIAL

Uso da inteligência artificial em Oncologia: *Doctor in silico?*

Artificial intelligence in Oncology: Doctor in silico?

Mateus Trinconi Cunha¹, Gilberto de Castro Junior²

A definição de um diagnóstico, e o respectivo tratamento, é baseada no uso das informações coletadas junto ao paciente, o que no contexto de experiências anteriores, permite associar o padrão observado (sinais, sintomas, exames complementares) a algum padrão conhecido. Assim, o (já enorme) volume crescente de informações disponíveis decorrentes de análises moleculares, aumento da qualidade e segurança de exames de imagem e progressivas descobertas acerca do processo saúde-doença permitem que possamos definir com maior precisão a melhor conduta em cada caso. Apesar desse progresso, o processamento de dados e avaliação de pacientes ainda depende de seres humanos: cansáveis, sujeitos a erros, com eventual alto custo e longo tempo de treinamento.

Nas últimas décadas, tem-se explorado o emprego de matemática e estatística para auxiliar na tomada de decisões clínicas e melhor organização de informações. Uma das linhas que mais tem recebido atenção recentemente é a inteligência artificial (IA). Essa é uma expressão que engloba diferentes algoritmos capazes de “aprender” (ou serem “treinados”) a realizar tarefas de classificação ou regressão a partir de informações anteriores sem programação explícita. Exemplos desses métodos incluem árvores de decisão, regressão linear, redes Bayesianas e redes neurais. Apesar de extremamente complexas, essas modalidades consistem em operações simples que, de forma geral, poderiam ser feitas por humanos, mas que por sua quantidade, levariam anos para serem feitas de maneira analógica, porém de segundos a minutos para máquinas.

Dessa forma, tem-se estudado as aplicações desta tecnologia em Medicina. No campo da análise de imagens, por exemplo, a análise de características pré-definidas (análise de forma, intensidade e homogeneidade de tons da região de interesse, e transformações matemáticas desses valores), e/ou uso de redes neurais e técnicas de visão de computador, são factíveis através de IA. Por meio desses artifícios, é possível encontrar correlações não identificadas por humanos.

Na patologia digital, as lâminas são digitalizadas com alta definição, sendo um pixel da imagem correspondente a até 0,25 micrômetros, e cada lâmina tendo em torno de 15 gigabytes de informação¹. A avaliação de diferentes características do tecido, desde formato celular a padrões de cor em núcleos e outras estruturas celulares, pode fornecer dados para identificação, por exemplo, de tecido estromal e epitelial,

1. Médico Residente do Serviço de Oncologia Clínica do Instituto do Câncer do Estado de São Paulo, Hospital das Clínicas da FMUSP, São Paulo, Brasil. <https://orcid.org/0000-0001-9101-8553>. E-mail: mateuscunha@gmail.com.
2. Médico Assistente do Serviço de Oncologia Clínica do Instituto do Câncer do Estado de São Paulo, Hospital das Clínicas da FMUSP, São Paulo, Brasil. Professor Doutor da Disciplina de Oncologia da FMUSP. <https://orcid.org/0000-0001-8765-3044>. E-mail: gilberto.castro@usp.br.

agrupamentos de células normais ou tumorais e até linfócitos infiltrantes de estroma ou tumor. Esse tipo de avaliação é denominado *Pathomics*. Essa tecnologia foi utilizada por Wang et al.², a fim de detectar um subgrupo de maior risco de recorrência em pacientes com câncer de pulmão não-células pequenas (CPCNP) em estádios I-II, sugerindo benefício em sobrevida livre de doença e sobrevida global com uso de quimioterapia adjuvante nessa população.

Uma estratégia similar de avaliação também é empregada no campo da Radiologia. Em um arquivo de exame de imagem, chamado DICOM, cada pixel pode ter de 12 a 16 bits, isso é, pode assumir de 4.096 a 65.536 tons de cinza possíveis. Os monitores usados para leitura de exames de imagem geralmente são capazes de processar até 8 bits por pixel, podendo assumir 256 tons de cinza³. Dessa forma, somente 0,39 a 6,25% da informação disponível chega aos olhos do Médico Radiologista. Assim, pode-se usar *Radiomics*, isso é, a análise de toda a informação bruta no arquivo, ajudando a a criar hipóteses e detectar padrões ainda não encontrados por humanos. Wang et al.⁴, por exemplo, ao combinar a avaliação de imagens feitas através de redes neurais, extração de variáveis pré-planejadas do tumor e características clínicas, conseguiu prever a presença de mutações no gene *EGFR* (do inglês *Epidermal Growth Factor Receptor*) e padrões de imunexpressão de PD-L1 em CPCNP a partir de imagens de tomografias computadorizadas de tórax com acurácia razoável (0,62-0,84). Essas variáveis são cotidianamente utilizadas na prática clínica na definição de tratamento do CPCNP.

Além do reconhecimento de imagens, a inteligência artificial também tem sido utilizada para a leitura de textos e extração de informações, na forma de processamento de linguagem natural (NLP, do inglês, *Natural Language Processing*). Essa modalidade é capaz de auxiliar pesquisadores a ler textos livres em prontuários e transformá-los em dados tabulares. Utilizando essa tecnologia, foi possível extrair dados de milhares de prontuários e compor classificações de risco para trombose venosa profunda recorrente e hemorragia em pacientes sob anticoagulação, com dados de mundo real^{5,6}.

Apesar de fornecerem conceitos interessantes e potencialmente úteis, para sua implantação na prática clínica, esses modelos devem passar por diversos passos, sendo eles: preparo e coleta de dados, criação do modelo inicial, validação externa, criação da plataforma de uso, avaliação do impacto na prática clínica e implementação⁷. A grande maioria dos modelos atualmente publicados ainda se encontram na fase de validação. Entretanto, isso não impede que a mídia leiga – e muitas vezes especializada – reporte os resultados como uma verdadeira revolução das máquinas, enquanto a realidade é que ainda estamos descobrindo os potenciais e limitações dessas técnicas, e as expectativas nelas colocadas podem não corresponder aos resultados^{8,9}.

Assim, ao se ler uma publicação sobre modelos de IA, deve-se compreender alguns dos vieses que podem estar presentes no estudo. A geração de um modelo confiável depende da entrada de grande quantidade de pacientes e de suas características. Subgrupos dessas características serão selecionados para que, em combinação, criem um valor de predição do desfecho estudado, sendo esse processo dependente de grande trabalho de geração de dados de qualidade, assim como de poder computacional adequado. Quanto à aplicabilidade externa do modelo desenvolvido, essa pode ser prejudicada por variáveis imprevistas inicialmente: por exemplo, variações de preparo de cortes histológicos ou execução de exames de imagem geram vieses de análise, ou seja, resultados errôneos decorrentes de dados de má qualidade, conhecido como “*garbage in – garbage out*”. Problemas de separação de dados podem também gerar correlações espúrias, que atrapalhariam a generalização para algumas populações. Finalmente, na produção de um resultado, muitas vezes não é possível compreender bem o que foi analisado pelo algoritmo, gerando um efeito de caixa preta, e incerteza na interpretação do dado obtido. Essas dificuldades, quando somadas, têm gerado grandes discussões no campo da IA na Medicina, como os possíveis vieses gerados na capacidade de identificação de etnia e sexo de pacientes através de exames de imagem, e na diferença de performance de algoritmos de acordo com essas variáveis^{10,11}. Para melhor identificar alguns desses vieses, diretrizes e critérios de qualidade têm sido estabelecidos nessa área⁷. Tecnologia diagnóstica através de algoritmo empregado na identificação de defeitos na recombinação homóloga do DNA, a partir de polimorfismos de nucleotídeo único (SNPs, *single nucleotide polymorphisms*), foi aprovada para uso clínico a fim de selecionar

aquelas pacientes candidatas ao uso de inibidores de PARP (*poly ADP-ribose polymerase*) como tratamento de manutenção em câncer de ovário¹².

Portanto, apesar do grande potencial da IA em desvendar padrões não identificáveis aos olhos humanos, ainda estamos longe de criar modelos complexos e seguros de qualidade comercial. A maior compreensão dessas técnicas e suas limitações possibilitarão a criação de algoritmos com capacidade de amplificar a acurácia e a produtividade dos médicos do futuro.

REFERÊNCIAS

1. Sakamoto T, Furukawa T, Lami K, Pham HHN, Uegami W, Kuroda K, et al. A narrative review of digital pathology and artificial intelligence: focusing on lung cancer. *Transl Lung Cancer Res.* 2020;9(5):2255–76. doi: <http://dx.doi.org/10.21037/tlcr-20-591>
2. Wang X, Bera K, Barrera C, Zhou Y, Lu C, Vaidya P, et al. A prognostic and predictive computational pathology image signature for added benefit of adjuvant chemotherapy in early stage non-small-cell lung cancer. *EBioMedicine.* 2021;69:103481. doi: <http://dx.doi.org/10.1016/j.ebiom.2021.103481>
3. Kimpe T, Tuytschaever T. Increasing the number of gray shades in medical display systems--how much is enough? *J Digit Imaging.* 2007;20(4):422–32. doi: <http://dx.doi.org/10.1007/s10278-006-1052-3>
4. Wang C, Ma J, Shao J, Zhang S, Liu Z, Yu Y, et al. Predicting EGFR and PD-L1 status in NSCLC patients using multitask AI system based on CT images. *Front Immunol.* 2022;13:813072. doi: <http://dx.doi.org/10.3389/fimmu.2022.813072>
5. Muñoz Martín AJ, Huerga Domínguez S, Souto JC, Rogado Revuelta J, Sanchez A, Garcia-Palomo A, et al. Predicting recurrence of venous thromboembolism in anticoagulated cancer patients using real-world data and machine learning. *J Clin Oncol.* 2022;40(16_suppl):e18742–e18742. doi: http://dx.doi.org/10.1200/jco.2022.40.16_suppl.e18742
6. Muñoz Martín AJ, Palacios ML, Souto JC, Obispo B, Aparicio J, Garcia-Palomo A, et al. Predicting major bleeding events in anticoagulated cancer patients with venous thromboembolism using real-world data and machine learning. *J Clin Oncol.* 2022;40(16_suppl):e18744–e18744. doi: http://dx.doi.org/10.1200/jco.2022.40.16_suppl.e18744
7. de Hond AAH, Leeuwenberg AM, Hooft L, Kant IMJ, Nijman SWJ, van Os HJA, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. *NPJ Digit Med.* 2022;5(1):2. doi: <http://dx.doi.org/10.1038/s41746-021-00549-7>
8. Dunnmon J. Separating hope from hype: Artificial intelligence pitfalls and challenges in radiology. *Radiol Clin North Am.* 2021;59(6):1063-74. doi: <http://dx.doi.org/10.1016/j.rcl.2021.07.006>
9. van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. *Nat Med.* 2021;27(5):775-84. doi: <http://dx.doi.org/10.1038/s41591-021-01343-4>
10. Gichoya JW, Banerjee I, Bhimireddy AR, Burns JL, Celi LA, Chen L-C, et al. AI recognition of patient race in medical imaging: a modelling study. *Lancet Digit Health.* 2022;4(6):e406-14. doi: [http://dx.doi.org/10.1016/S2589-7500\(22\)00063-2](http://dx.doi.org/10.1016/S2589-7500(22)00063-2)
11. Cirillo D, Catuara-Solarz S, Morey C, Guney E, Subirats L, Mellino S, et al. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. *NPJ Digit Med.* 2020;3(1):81. doi: <http://dx.doi.org/10.1038/s41746-020-0288-5>
12. myChoice CDx® Technical Information [cited July 25, 2022]. Available from: https://www.accessdata.fda.gov/cdrh_docs/pdf19/P190014S003C.pdf