



O PAPEL EMERGENTE DA INTELIGÊNCIA ARTIFICIAL (IA) E MACHINE LEARNING (ML) NO DIAGNÓSTICO E PROGNÓSTICO DAS NEOPLASIAS MIELOPROLIFERATIVAS (NMP)

ANA LUIZA VIEIRA; NATÁLIA PEROTTONI; LAÍS ZOTTE; LUCAS PASSOS DOS SANTOS; BEATRIZ ESTRELA

RESUMO

Introdução: As Neoplasias Mieloproliferativas (NMP), incluindo Policitemia Vera, Trombocitemia Essencial e Mielofibrose, são doenças hematológicas crônicas marcadas por heterogeneidade clínica e molecular. Essa complexidade dificulta o diagnóstico precoce e a estratificação prognóstica, tornando-se um desafio para os métodos convencionais. Nesse cenário, a Inteligência Artificial (IA) e o Machine Learning (ML) emergem como tecnologias promissoras na análise de dados clínicos, genéticos e laboratoriais, com potencial de aprimorar significativamente a acurácia diagnóstica e a predição de desfechos. **Objetivo:** Investigar o papel da IA e do ML no aprimoramento do diagnóstico, prognóstico e identificação de biomarcadores nas NMP. **Metodologia:** Trata-se de uma revisão narrativa baseada em artigos originais, revisões e diretrizes publicadas entre 2020 e 2025, nas línguas português, inglês e espanhol. A estratégia de busca utilizou termos MeSH e palavras-chave conforme a estrutura PICO. Foram incluídos estudos que abordam a aplicação de IA e ML em NMPs, sendo excluídos relatos de caso, estudos em animais e resumos sem texto completo. **Resultados:** Diversos modelos baseados em IA demonstraram desempenho diagnóstico comparável ou superior ao de especialistas, utilizando dados hematológicos, transcriptômicos e genômicos. Algoritmos como redes neurais convolucionais, XGBoost e GMNB alcançaram altas taxas de acurácia na distinção entre subtipos de NMP e na previsão de mutações. Além disso, sistemas preditivos como o AIPSS-MF superaram os modelos prognósticos tradicionais, mesmo sem dados genéticos. A IA também se mostrou eficaz na identificação de novos biomarcadores genéticos e proteômicos, com aplicação potencial na personalização terapêutica. **Conclusão:** A integração de IA e ML na prática hematológica representa um avanço significativo no diagnóstico e prognóstico das NMPs, oferecendo maior precisão e padronização. No entanto, sua adoção clínica exige validação robusta, padronização de dados e superação de barreiras éticas e interpretativas. Estas ferramentas devem atuar como suporte ao julgamento clínico, promovendo uma medicina mais personalizada e eficiente.

Palavras-chave: Neoplasias Mieloproliferativas; Inteligência Artificial; Machine Learning

1 INTRODUÇÃO

As neoplasias mieloproliferativas (NMPs), representadas principalmente por Policitemia Vera, Trombocitemia Essencial e Mielofibrose, constituem um grupo heterogêneo de doenças hematológicas que apresentam desafios significativos em relação ao diagnóstico preciso e à estratificação de risco, devido à diversidade clínica e molecular de cada condição (Han et al., 2024). Métodos diagnósticos tradicionais têm se mostrado limitados diante da complexidade dessas patologias, frequentemente resultando em incertezas clínicas que impactam diretamente as estratégias terapêuticas adotadas (Sampaio et al., 2023). Assim, surge a necessidade urgente por abordagens inovadoras capazes de integrar eficientemente dados clínicos, genéticos e imunohistoquímicos, garantindo maior precisão na identificação,

classificação e manejo dos pacientes com essas enfermidades (Cardoso et al., 2024).

Nesse contexto, a utilização de Inteligência Artificial (IA) e Machine Learning (ML) vem despontando como uma solução promissora para superar tais limitações. Essas tecnologias possibilitam não apenas a análise de grandes volumes de dados complexos e heterogêneos, mas também oferecem potencial para aprimorar substancialmente a acurácia diagnóstica e preditiva em NMPs. Evidências recentes sugerem que modelos baseados em IA podem alcançar precisão diagnóstica superior a 90%, destacando-se especialmente na identificação precoce da doença e na predição personalizada dos desfechos clínicos, superando frequentemente as técnicas convencionais (Sampaio et al., 2023; Visconte et al., 2024). Contudo, apesar desse potencial transformador, a implementação clínica desses modelos ainda enfrenta importantes desafios técnicos, éticos e regulatórios, que precisam ser devidamente compreendidos e superados para viabilizar seu uso rotineiro na prática médica (Cardoso et al., 2024).

Diante disso, o objetivo desta revisão é analisar criticamente os avanços recentes na aplicação da Inteligência Artificial e do Machine Learning para o diagnóstico e a estratificação prognóstica das neoplasias mieloproliferativas, abordando suas vantagens, limitações e perspectivas futuras, com o intuito de identificar lacunas científicas e clínicas que possam orientar futuras pesquisas e aplicações clínicas nesta área emergente (Han et al., 2024; Visconte et al., 2024).

2 MATERIAL E MÉTODOS

A metodologia adotada para esta revisão narrativa foi elaborada com o objetivo de explorar o papel crescente da Inteligência Artificial e do Machine Learning no diagnóstico e prognóstico das Neoplasias Mieloproliferativas. Para isso, realizou-se uma busca abrangente de literatura científica, utilizando termos padronizados do vocabulário MeSH e palavras-chave relacionadas, de modo a garantir a seleção de estudos relevantes. A elaboração da estratégia de busca seguiu a estrutura PICO, sendo que a população de interesse incluiu pacientes com Neoplasias Mieloproliferativas, como Policitemia Vera, Trombocitemia Essencial e Mielofibrose. A intervenção considerada foi o uso de algoritmos de Inteligência Artificial e Machine Learning aplicados ao contexto clínico dessas doenças. As abordagens convencionais de diagnóstico e prognóstico serviram como comparação, incluindo critérios morfológicos, histopatológicos e moleculares tradicionais. Os desfechos avaliados englobaram o aprimoramento da acurácia diagnóstica, a estratificação de risco, a previsão da progressão da doença e a descoberta de novos biomarcadores prognósticos.

A busca bibliográfica utilizou uma combinação específica de termos relacionados às neoplasias mieloproliferativas, inteligência artificial e avaliação diagnóstica e prognóstica. Foram consultadas bases de dados científicas relevantes e incluídos estudos publicados entre os anos de 2020 e 2025, redigidos nos idiomas português, inglês ou espanhol. Foram considerados elegíveis artigos originais, revisões e diretrizes. Foram excluídos estudos realizados exclusivamente em animais, relatos de casos isolados e publicações sem acesso ao texto completo. Esta abordagem metodológica visou reunir evidências atualizadas e representativas sobre a integração da inteligência artificial nas práticas clínicas relacionadas às neoplasias mieloproliferativas.

3 RESULTADOS E DISCUSSÃO

A aplicação de algoritmos de Inteligência Artificial (IA) e Machine Learning (ML) no diagnóstico diferencial das Neoplasias Mieloproliferativas (NMP) tem se mostrado promissora, especialmente para casos Philadelphia-negativos, nos quais há sobreposição de características entre subtipos como Policitemia Vera, Trombocitemia Essencial e Mielofibrose Primária. Kimura et al. (2021) demonstraram que um sistema automatizado, integrando um analisador hematológico com uma rede neural convolucional e o modelo XGBoost, foi capaz de distinguir

esses subtipos com alta acurácia (AUCs superiores a 0,96), utilizando apenas dados de sangue periférico. A interpretação por SHAP contribuiu para maior transparência das decisões. Em paralelo, Zhang et al. (2022) desenvolveram um modelo baseado em Geometric Mean Naïve Bayes (GMNB), treinado com dados transcriptômicos de mais de 4.600 amostras, que distinguiu 45 entidades hematológicas, incluindo NMPs, com AUCs robustos, como 0,925 para diferenciação entre MPN e controles normais. Apesar da acurácia reduzida quando se considerava apenas a primeira hipótese diagnóstica, a performance aumentou ao incluir a segunda opção, reforçando o papel da IA como suporte complementar à decisão clínica e à triagem diagnóstica, especialmente em contextos com recursos limitados ou sobreposição fenotípica.

Esses achados indicam que algoritmos de IA baseados em transcriptoma podem atuar como filtros iniciais poderosos, capazes de orientar exames complementares com maior especificidade, acelerando a definição diagnóstica e reduzindo erros associados à variabilidade interobservadora. Avançando nessa linha, Zhang et al. (2023) aprofundaram a aplicação de algoritmos de IA e ML para o diagnóstico diferencial das neoplasias hematológicas, utilizando uma abordagem baseada em transcriptoma direcionado (sequenciamento de RNA de 1408 genes). Em um conjunto de 88 casos de NMP, o modelo (Geometric Mean Naïve Bayes) diferenciou NMP de MDS com sensibilidade de 90,9% e especificidade de 70,8%, mesmo sem dados mutacionais. Esses resultados reforçam o valor do perfil transcriptômico associado à IA na identificação de padrões diferenciais de expressão gênica entre entidades mieloides sobrepostas, desafiando critérios morfológicos convencionais que falham frente à subjetividade ou limitações amostrais (ZHANG, 2023). O modelo também se mostrou eficaz na distinção entre amostras normais de medula óssea e NMP (AUC = 0,923), útil no diagnóstico inicial. Embora o desempenho na distinção entre subtipos específicos (PV e TE) não tenha sido detalhado, a arquitetura do modelo permite tal aplicação. A acurácia para NMP como primeira escolha em múltiplas classes ainda é limitada (12%), revelando a necessidade de integração com dados clínicos e moleculares complementares (mutações JAK2, CALR ou MPL) para aumento da acurácia diagnóstica final. Portanto, no contexto das NMP, algoritmos baseados em IA aliados à expressão gênica dirigida podem melhorar substancialmente o diagnóstico diferencial inicial, atuando como ferramenta auxiliar para patologistas, reduzindo a subjetividade e oferecendo diagnósticos mais padronizados.

Além do diagnóstico, algoritmos de machine learning têm revolucionado a classificação e o prognóstico das Neoplasias Mieloproliferativas (NMP), ao integrar dados clínicos e genômicos com alta precisão. Grinfeld et al. (2018) analisaram mais de 2000 pacientes com NMP, identificando genes condutores e construindo modelos preditivos baseados em 63 variáveis, capazes de antecipar progressão para mielofibrose, transformação leucêmica e sobrevida livre de eventos com desempenho superior aos sistemas prognósticos tradicionais. Mutações específicas em genes epigenéticos, de splicing e da via RAS foram associadas a piores desfechos, reforçando o papel da genômica expandida na estratificação de risco. Esses modelos, além de precisos, oferecem curvas de sobrevida personalizadas, o que exemplifica o potencial da IA para orientar terapias individualizadas. No caso específico da mielofibrose, Duminuco et al. (2023) desenvolveram o AIPSS-MF, um modelo preditivo treinado com dados clínicos de mais de 1300 pacientes, capaz de estimar sobrevida global e livre de leucemia sem necessidade de informações genômicas. O AIPSS-MF mostrou desempenho superior aos sistemas convencionais e ampliou o acesso à estratificação individualizada, mesmo em contextos com recursos limitados, demonstrando como a IA pode democratizar a medicina de precisão em NMP.

A aplicação de algoritmos de IA e ML tem demonstrado notável potencial na descoberta e validação de biomarcadores em diversas neoplasias hematológicas, incluindo as NMP. Essas ferramentas computacionais, ao explorarem grandes volumes de dados clínicos, genômicos e

proteômicos, superam limitações dos métodos tradicionais, permitindo a identificação de padrões complexos. Nesse sentido, Cheng et al. (2024) aplicaram diferentes algoritmos de ML (LASSO, SVM-RFE, Random Forest e XGBoost) para triagem e validação de biomarcadores genéticos com valor prognóstico em pacientes com leucemia mieloide aguda (LMA), condição relacionada às NMP. Eles identificaram três genes (DNM1, MEIS1 e SUSD3) com relevância prognóstica, confirmando a robustez dos modelos por validação cruzada e análise pan-câncer (CHENG, 2024). Esses resultados reforçam a aplicabilidade desses métodos também para NMP, especialmente para previsão de transformação leucêmica e estratificação de risco.

No contexto proteômico, Mann et al. (2021) destacam o avanço do uso da IA na análise de espectrometria de massas em larga escala. Algoritmos de deep learning, como redes neurais convolucionais (RNCs) e recorrentes (RNRs), possibilitaram a previsão de espectros de fragmentação peptídica com altíssima acurácia, um marco para a identificação de biomarcadores proteômicos, pois aumentam a sensibilidade e especificidade da detecção de proteínas diferencialmente expressas em doenças hematológicas. Estratégias como "transfer learning" permitem adaptar modelos preexistentes, reduzindo a dependência de grandes volumes de dados (MANN, 2021). A capacidade de integrar dados proteômicos com outras camadas ômicas potencializa a identificação de assinaturas moleculares mais completas e clinicamente relevantes para as NMP. Adicionalmente, o trabalho de Kewan et al. (2023) demonstrou a eficácia de técnicas de clustering não supervisionado em coortes de pacientes com neoplasias mieloides, incluindo NMP e síndromes mielodisplásicas (SMD). A análise identificou nove grupos moleculares distintos com perfis genéticos e prognósticos específicos, revelando subgrupos previamente não reconhecidos pelas classificações morfológicas convencionais e fornecendo bases para a personalização terapêutica (KEWAN, 2023).

A análise de dados oriundos de tecnologias ômicas, como o sequenciamento de nova geração (NGS) e a citometria de fluxo, apresenta desafios substanciais na prática clínica, especialmente em NMP, caracterizadas por heterogeneidade genética e evolução clonal complexa. Nesse contexto, a aplicação de IA e ML surge como ferramenta promissora para lidar com o alto volume, a variabilidade e a complexidade dos dados gerados. O estudo de Park et al. (2021), por exemplo, oferece uma revisão aprofundada sobre a integração de dados NGS com algoritmos de IA, destacando o uso de arquiteturas de deep learning (RNCs e GNNs). Essas arquiteturas correlacionam mutações genéticas com perfis de expressão gênica, modificações epigenéticas e dados clínicos, permitindo a identificação de padrões moleculares não evidentes com métodos estatísticos tradicionais. O artigo reforça que essas abordagens são particularmente úteis em condições hematológicas, como as NMP, nas quais mutações como JAK2, CALR e MPL coexistem com outras alterações secundárias que influenciam a evolução da doença (PARK, 2021).

De forma complementar, Schmidt e Hildebrandt (2021) analisam os avanços no uso de deep learning especificamente em dados de NGS, com ênfase na chamada de variantes, predição funcional de mutações e integração com dados transcriptômicos. Os autores destacam que técnicas como redes neurais profundas (DNN) e autoencoders são capazes de realizar tarefas como normalização de ruído técnico e detecção de padrões genômicos complexos sem extração manual de características. Tais estratégias aumentam significativamente a sensibilidade na identificação de mutações, com impacto direto na acurácia diagnóstica e prognóstica, superando a subjetividade e as limitações dos pipelines bioinformáticos convencionais (SCHMIDT, 2021). Já o estudo aplicado de Pellegrino et al. (2021) exemplifica o uso prático de ML na análise de dados NGS oncosomáticos. O grupo implementou um modelo baseado em Random Forest (RF) para classificar variantes genéticas obtidas por NGS, a partir de um painel de 27 genes tumorais. O algoritmo alcançou precisão superior a 99% na diferenciação entre variantes patogênicas e benignas (PELLEGRINO, 2021). Embora o estudo envolva cânceres sólidos, os princípios metodológicos são totalmente transponíveis para as

NMP, onde a classificação de mutações somáticas e o reconhecimento de padrões genéticos têm papel central.

A aplicação de algoritmos de Inteligência Artificial (IA) e Aprendizado de Máquina (AM) tem avançado no contexto das Neoplasias Mieloproliferativas (NMP), com destaque para o diagnóstico, previsão de mutações e estratificação de risco. Estudo de Choi et al. (2024) demonstrou que redes neurais convolucionais (RNCs), como a EfficientNet-B1, superam SVMs e árvores de decisão na previsão de mutações em genes como JAK2 a partir de imagens de esfregaço sanguíneo, com AUC de 0,87. As RNCs se destacam por extraírem automaticamente padrões visuais, dispensando processamento manual, enquanto SVMs e Random Forests são mais interpretáveis e menos exigentes computacionalmente, embora com menor desempenho preditivo. No entanto, as RNCs apresentam limitações como a falta de transparência decisória (“caixa preta”) e a exigência de grandes volumes de dados anotados — um desafio para doenças raras como as NMPs. Como destacam Gomes e Ashley (2023), o diagnóstico dessas neoplasias ainda é subjetivo e dependente de critérios morfológicos e moleculares, e a IA pode contribuir para maior objetividade e padronização, especialmente se integrada a dados clínicos e laboratoriais. Apesar do potencial, a aceitação clínica ainda depende de validação rigorosa, padronização dos dados e superação de barreiras relacionadas à interpretabilidade e reprodutibilidade dos modelos.

4 CONCLUSÃO

A Inteligência Artificial (IA) e o Machine Learning (ML) representam um avanço promissor no campo das Neoplasias Mieloproliferativas (NMPs), complementando as abordagens diagnósticas e prognósticas tradicionais. Essas tecnologias aprimoram o diagnóstico ao identificar padrões complexos e reconhecer subtipos atípicos, enquanto no prognóstico, oferecem uma estratificação de risco mais precisa pela incorporação de variáveis não lineares e atualização dinâmica de previsões. Em um escopo translacional, a IA/ML facilita a análise integrada de dados ômicos, genéticos e clínicos, permitindo a descoberta de novos biomarcadores moleculares e epigenéticos. Essa capacidade de processar dados multimodais em larga escala acelera o raciocínio clínico e impulsiona uma medicina mais personalizada.

REFERÊNCIAS

CHENG, Y. et al. Multiple machine-learning tools identifying prognostic biomarkers for acute Myeloid Leukemia. **BMC Medical Informatics and Decision Making**, [S. l.], v. 24, n. 1, p. 2, 2 jan. 2024. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/38167056/>. Acesso em: 18 jul. 2025.

CHOI, J.; LEE, H.; KIM-GODWIN, Y. Decoding machine learning in nursing research: A scoping review of effective algorithms. **Journal of Nursing Scholarship**, [S. l.], v. 57, n. 1, p. 119-129, jan. 2025. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/39294553/>. Acesso em: 18 jul. 2025.

DUMINUCO, A. et al. Myelofibrosis and Survival Prognostic Models: A Journey between Past and Future. **Journal of Clinical Medicine**, [S. l.], v. 12, n. 6, p. 2188, 11 mar. 2023. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10053868/>. Acesso em: 18 jul. 2025.

GOMES, B.; ASHLEY, E. A. Artificial intelligence in molecular medicine. **New England Journal of Medicine**, [S. l.], v. 388, n. 26, p. 2510-2521, 29 jun. 2023. Disponível em: <https://www.nejm.org/doi/full/10.1056/NEJMra2204787>. Acesso em: 18 jul. 2025.

GRINFELD, J. et al. Classification and Personalized Prognosis in Myeloproliferative Neoplasms. **New England Journal of Medicine**, [S. l.], v. 379, n. 15, p. 1416-1430, 2018.

HAN, J. et al. Applications of Artificial Intelligence in Philadelphia-Negative Myeloproliferative Neoplasms. In: **Advancements in Machine Learning (ML): Transforming the Future of Blood Cancer Detection and Outcome Prediction**. [S. l.]: healthbook TIMES Oncology Hematology, 2024.

KEWAN, T. et al. Molecular patterns identify distinct subclasses of myeloid neoplasia. **Nature Communications**, [S. l.], v. 14, n. 1, p. 3136, 30 maio 2023. Erratum in: *Nat Commun.* 2024 Feb 8;15(1):1184. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37253784/>. Acesso em: 18 jul. 2025.

KIMURA, K. et al. Automated diagnostic support system with deep learning algorithms for distinction of Philadelphia chromosome-negative myeloproliferative neoplasms using peripheral blood specimen. **Scientific Reports**, [S. l.], v. 11, n. 1, p. 3367, 9 fev. 2021.

MANN, M.; KUMAR, C.; ZENG, W. F.; STRAUSS, M. T. Artificial intelligence for proteomics and biomarker discovery. **Cell Systems**, [S. l.], v. 12, n. 8, p. 759-770, 18 ago. 2021. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/34411543/>. Acesso em: 18 jul. 2025.

PARK, Y.; HEIDER, D.; HAUSCHILD, A. C. Integrative Analysis of Next-Generation Sequencing for Next-Generation Cancer Research toward Artificial Intelligence. **Cancers**, [S. l.], v. 13, n. 13, p. 3148, 24 jun. 2021. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8269018/>. Acesso em: 18 jul. 2025.

PELLEGRINO, E. et al. Machine learning random forest for predicting oncosomatic variant NGS analysis. **Scientific Reports**, [S. l.], v. 11, n. 1, p. 21820, 8 nov. 2021. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/34750410/>. Acesso em: 18 jul. 2025.

SCHMIDT, B.; HILDEBRANDT, A. Deep learning in next-generation sequencing. **Drug Discovery Today**, [S. l.], v. 26, n. 1, p. 173-180, jan. 2021. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/33059075/>. Acesso em: 18 jul. 2025.

ZHANG, H. et al. Differential diagnosis of hematologic and solid tumors using targeted transcriptome and artificial intelligence. **Journal of Clinical Oncology**, [S. l.], v. 40, n. 16, p. 3018, 2022. Suplemento 16. Disponível em: <https://meetings.asco.org/abstracts-presentations/207880>. Acesso em: 18 jul. 2025.

ZHANG, H. et al. Differential Diagnosis of Hematologic and Solid Tumors Using Targeted Transcriptome and Artificial Intelligence. **The American Journal of Pathology**, [S. l.], v. 193, n. 1, p. 51-59, 2023. Disponível em: <https://ajp.amjpathol.org/action/showPdf?pii=S0002-9440%2822%2900313-3>. Acesso em: 18 jul. 2025.