



Desenvolvimento de Modelos Preditivos para o Prognóstico da Evolução Clínica de Pacientes com COVID-19 a partir do Ar Exalado

Arlene B. S. Nossol (PQ)^{1*}, Débora L. Z. Scheucher (PQ)², Luciana M. Bastos (PQ)¹, Mário M. Martins (PQ)¹, Thulio M. Cunha (PQ)²

¹ Laboratório de Nanobiotecnologia Prof. Dr. Luiz Ricardo Goulart Filho, Instituto de Biotecnologia, Universidade Federal de Uberlândia, Uberlândia, 38405-320, Brasil.

² Faculdade de Medicina Universidade Federal de Uberlândia, Uberlândia, 38400-902, Brasil.

E-mail: arlene.nossol@ufu.br

RESUMO

A análise de compostos orgânicos voláteis (COVs) no ar exalado tem se consolidado como uma estratégia promissora para o monitoramento da evolução da COVID-19. Este estudo investigou os perfis metabólicos de pacientes infectados pelo SARS-CoV-2 utilizando cromatografia gasosa acoplada à espectrometria de massas (CG-EM), identificando 10 metabólitos voláteis de maior relevância para a classificação entre os grupos "alta" e "intubado", com destaque para o dicloroacetaldeído, que apresentou o maior peso na diferenciação entre os grupos. Para a classificação dos grupos, o melhor modelo preditivo foi o Gradient Boosting, que alcançou uma sensibilidade de 80,0%, especificidade de 72,7% e acurácia de 76,9% na distinção entre pacientes que receberam alta e aqueles que necessitaram de intubação. Os resultados indicam que os COVs refletem alterações metabólicas, inflamatórias e de estresse oxidativo, posicionando-se como potenciais biomarcadores não invasivos para suporte clínico em ambientes hospitalares.

Palavras-chave: COVID-19, compostos orgânicos voláteis, CG-EM, modelos preditivos, biomarcadores

Introdução

A COVID-19, causada pelo vírus SARS-CoV-2, representa um dos maiores desafios sanitários contemporâneos (1, 2). A identificação de biomarcadores não invasivos para o prognóstico da doença é fundamental para auxiliar no manejo clínico dos pacientes (3, 4). Compostos orgânicos voláteis (COVs) presentes no ar exalado podem refletir alterações metabólicas associadas à inflamação e ao estresse oxidativo, sendo esses metabólitos potenciais biomarcadores para a detecção precoce da doença (5, 6, 7). Neste contexto, a análise do ar exalado surge como uma alternativa promissora, especialmente em ambientes com recursos limitados, pois oferece um método simples e não invasivo para o prognóstico da evolução da COVID-19 (8, 9).

Experimental

Análise dos metabólitos.

Amostras de ar exalado foram coletadas de 26 pacientes internados (11 pacientes do grupo alta x 15 pacientes do grupo intubado) com diagnóstico confirmado de COVID-19, em períodos pré-determinados, no Hospital de Clínicas da Universidade Federal de Uberlândia (HC-UFU). Este estudo foi aprovado pelo Comitê de Ética em Pesquisa com Seres Humanos, sob o número do CAAE: 30848620.1.3009.5152.

As análises dos metabólitos presentes nas amostras foram realizadas utilizando cromatógrafo a gás acoplado à espectrômetro de massas (CG-EM) modelo Agilent 7890B GC/5977B MSD, foi utilizada a coluna cromatográfica capilar Agilent 19091S-433 HP-5ms (30 m x 250 μ m x 0.25 μ m). As condições cromatográficas empregadas foram: temperatura do injetor, 210 °C. O gás de arraste era hélio de alta pureza a uma taxa

de fluxo constante de 1,0 mL min⁻¹. Foram feitas injeções de 2 mL utilizando o modo Split 1:1. O modo Scan (MS) foi configurado para m/z 35–550, com uma contagem de 150 limiar. O quadrupolo foi mantido a 150 °C com a fonte de íons a 280 °C. A temperatura inicial do forno foi programada de 40 °C aumentando a uma taxa de 3°C min⁻¹ até 220 °C, mantendo a essa temperatura por 60 minutos. A temperatura de incubação foi de 150 °C, mantida por 6 minutos.

Análise estatística dos dados.

Os dados gerados durante as análises foram processados utilizando o software *MassHunter Unknowns Analysis 10.0* com os espectros convertidos para o formato .CEF e criação da tabela de dados pelo *Mass Profiler Professional (MPP)*. Modelos preditivos foram então gerados por meio de algoritmos do software *Orange 3.37.0*, e treinados para gerar as estatísticas relacionadas à sensibilidade, especificidade e acurácia dos modelos foram avaliadas utilizando a plataforma online MEDCALC.

Resultados e Discussão

A coleta de amostras de ar exalado foi realizada em 26 pacientes diagnosticados com COVID-19, e as análises dos metabólitos foram realizadas por CG-EM. Os cromatogramas foram processados, gerando uma tabela de dados no MPP. Os dados foram então normalizados e divididos para treinamento utilizando a técnica de validação cruzada *leave-one-out* (LOO). Diversos modelos preditivos (Random Forest, Gradient



Boosting, SVM, Neural Network, Decision Tree, Logistic Regression, AdaBoost, Naive Bayes) foram aplicados e avaliados no *software Orange 3.37.0*. O modelo Gradient Boosting apresentou o melhor desempenho, classificando com maior precisão os pacientes em dois grupos: alta e intubação. A matriz de confusão permitiu uma clara diferenciação entre os dois grupos (Fig. 1). A Figura 1 exibe o número de pacientes corretamente classificados: 8 no grupo de alta e 12 no grupo de intubação, além dos erros de classificação, com 3 pacientes do grupo de alta classificados como intubados e 3 pacientes do grupo de intubação classificados como alta.

		Predicted		
		Alta	Intubado	Σ
Actual	Alta	8	3	11
	Intubado	3	12	15
Σ		11	15	26

Figura 1. Matriz de Confusão mostrando a performance do modelo Gradient Boosting na classificação dos pacientes em dois grupos: alta e intubação.

Os resultados estatísticos revelaram a eficácia do modelo preditivo, com destaque para a sensibilidade de 80,0% indica boa capacidade do modelo em identificar corretamente os pacientes que necessitaram de intubação, enquanto a especificidade de 72,7% mostra que o modelo é razoavelmente eficaz para classificar os pacientes (Tabela 1).

Tabela 1. Parâmetros estatísticos do modelo preditivo Gradient Boosting, para a análise do ar exalado de pacientes com COVID-19.

Métrica	Valor (%)
Sensibilidade	80,0
Especificidade	72,7
Acurácia	76,9

A análise dos metabólitos voláteis presentes no ar exalado revelou 10 compostos de maior relevância para a classificação dos grupos alta versus intubado, com destaque para o dicloroacetaldeído, que apresentou o maior peso na diferenciação entre os grupos. Outros metabólitos identificados incluem: 2-cianoguanidina, N-metilmetacrilamida, acetaldeído, tetramero, 1,3-dioxol-2-ona, propano, 2-metil-2-nitro-, peróxido de dietila, 1H-tetrazol, 1H-pirazol-4-carbonitrila e N-benzyloxi-2-carbometoxila. A presença desses metabólitos foi associada a alterações metabólicas características da COVID-19, como estresse oxidativo e inflamação. O dicloroacetaldeído, em particular, apresenta grande potencial como biomarcador para o diagnóstico precoce e monitoramento da doença, oferecendo uma alternativa não invasiva às ferramentas diagnósticas convencionais. Esses resultados indicam que a análise do ar

exalado, aliada a métodos de aprendizado de máquina, possui grande potencial para complementar as ferramentas diagnósticas da COVID-19, especialmente em cenários de alta demanda ou recursos limitados.

Conclusões

Os resultados indicam que os COVs presentes no ar exalado de pacientes com COVID-19 têm grande potencial como biomarcadores para diagnóstico e monitoramento da doença. No entanto, a presença de falsos negativos e falsos positivos destaca a necessidade de aprimoramentos no modelo, como a inclusão de mais variáveis clínicas e a validação com conjuntos de dados maiores. A utilização de métodos de aprendizado de máquina demonstrou ser uma abordagem promissora para melhorar a tomada de decisão clínica, mas estudos adicionais são necessários para refinar os algoritmos e otimizar a precisão diagnóstica.

Agradecimentos

Aos profissionais e pacientes do Hospital de Clínicas pela colaboração essencial na coleta das amostras. Aos órgãos de fomento, como a FAPEMIG pelo apoio financeiro. Aos membros da equipe de pesquisa — Débora, Lucina Silva, Luciana Bastos, professor Thulio e Mário — pela assistência na coleta, orientação e valiosa contribuição técnica durante o desenvolvimento deste estudo. Agradecemos também aos laboratórios e centros de pesquisa envolvidos, como Laboratório de Nanobiotecnologia Prof. Dr. Luiz Ricardo Goulart Filho, pelo suporte e recursos fornecidos. Por fim, agradecemos especialmente aos participantes do estudo pela confiança e colaboração essenciais para a realização desta pesquisa.

Referências

- Corman, V. M., et al. "Detection of 2019-nCoV in saliva." *Lancet* (2020).
- Grassin-Delyle, S., et al. "Exhaled volatile organic compounds as biomarkers in respiratory disease." *Chest* (2021).
- Ruszkiewicz, J., et al. "Volatile organic compounds in breath as a diagnostic tool for COVID-19." *J. Clin. Microbiol.* (2020).
- Nigro, P., et al. "SARS-CoV-2 detection from exhaled air using ion mobility spectrometry." *Sci. Rep.* (2020).
- Ibrahim, H. M., et al. "Volatile organic compounds in breath as biomarkers of COVID-19." *J. Breath Res.* (2021).
- Grassin-Delyle, S., et al. "Breath analysis for medical diagnostics." *J. Clin. Med.* (2021).
- Woollam, R., et al. "Breath analysis in early diagnosis of COVID-19 infection." *Lancet Respir. Med.* (2022).
- Simonnet, L., et al. "Breath biomarkers for COVID-19 diagnosis and monitoring." *J. Med. Virol.* (2020).
- Shi, L., et al. "Detection of SARS-CoV-2 through VOCs in breath samples." *Anal. Chem.* (2020).