

ACURÁCIA DIAGNÓSTICA DA INTELIGÊNCIA ARTIFICIAL NA INTERPRETAÇÃO DE TRAÇADOS DE HOLTER EM CÃES: ESTUDO-PILOTO CONTROLADO

Pamela Regina Pimenta Busato¹

Jaqueline Yamane²

Matheus Felipe Souza Franco³

Susamara De Souza Da Silva⁴

Tatiana Champion⁵

Palavras-chave: Análise automatizada; Arritmias cardíacas; Eletrocardiografia ambulatorial; Algoritmo; Medicina Veterinária.

INTRODUÇÃO

O holter é um método pouco invasivo, considerado um valioso recurso diagnóstico na cardiologia por fornecer grandes informações sobre a atividade elétrica cardíaca, contribuindo na identificação de distúrbios do ritmo e condução elétrica do coração (Mavropoulou *et al.*, 2021; Lima *et al.*, 2025). A interpretação dos traçados eletrocardiográficos podem sofrer pequenas variações dependendo da experiência do profissional e acurácia na análise (Barros; Berwanger, 2025). Nos últimos anos a utilização de inteligência artificial (IA) aplicada no reconhecimento de traçados eletrocardiográficos tem se tornado mais frequente devido à agilidade no processamento de informações, principalmente em estudos que utilizam modelos de IA de alto aprendizado em ECG (Ose *et al.*, 2024).

Segundo Ose *et al.* (2024), IA's específicas para interpretação de eletrocardiograma tem se mostrado superior aos exames analisados por cardiologistas. Já em outro estudo realizado por Günay *et al.* (2024), interpretações de cardiologistas, superaram as análises realizadas por IA's (GPT- 4, GPT- 4.0 e Gemini).

¹ Universidade Federal da Fronteira Sul, pamsbusato@gmail.com

² Universidade Federal da Fronteira Sul, jaqueline.yamane@estudante.uffs.edu.br

³ Universidade Federal da Fronteira Sul, mfelipefranco@gmail.com

⁴ Universidade Federal da Fronteira Sul.

⁵ Universidade Federal da Fronteira Sul, tatiana.champion@uffs.edu.br

Há diversos desafios técnicos e regulamentações na incorporação no uso da IA na medicina, principalmente em cardiologia, que necessitam análises minuciosas e precisas, sendo necessários mais estudos sobre os limites e confiabilidade nesses algoritmos (Muzammil, *et al.* 2024; Barros & Berwanger, 2025). Dessa forma, o objetivo deste trabalho foi comparar a interpretação dos traçados eletrocardiográficos de Holter em cães obtida por IA, com a interpretação realizada por médicos veterinários atuantes na área de cardiologia, buscando avaliar se há concordância diagnóstica entre ambos, assim como a praticidade de uso da ferramenta. Por fim, pretendeu-se refletir sobre a aplicabilidade clínica da inteligência artificial como recurso auxiliar no diagnóstico cardiológico veterinário.

MATERIAIS E MÉTODOS

Foi conduzido um estudo-piloto transversal de acurácia diagnóstica para avaliar a concordância entre a interpretação de traçados de Holter por um sistema de inteligência artificial baseado em ChatGPT e o padrão-ouro de especialistas veterinários. O delineamento metodológico seguiu as recomendações STARD 2015 para estudos de acurácia diagnóstica.

O estudo foi desenvolvido no Serviço de Cardiologia Veterinária da Superintendência da Unidade Hospitalar Veterinária da Universidade Federal da Fronteira Sul, Campus Realeza, Paraná, Brasil, no período de agosto de 2024 a julho de 2025. Foram analisados 20 segmentos eletrocardiográficos de registros Holter de 24 horas de 7 cães, adquiridos com o sistema Cardios – CardioSmart 540 (modelo CardioLight; Cardios®, São Paulo, Brasil), padronizados em 25 mm/s, 10 mm/mV e três derivações simultâneas.

Os segmentos foram selecionados intencionalmente para incluir diferentes padrões eletrocardiográficos: ritmo sinusal, taquicardia sinusal, bradicardia sinusal, bem como a presença de complexos ventriculares prematuros (VPCs), complexos supraventriculares prematuros (SVPCs), bloqueios atrioventriculares (BAVs) dos tipos Mobitz I e II e pausas sinusais iguais ou superiores a dois segundos. A duração média foi de $6,1 \pm 2,0$ segundos (3,8–12,1s), com 7 a 22 batimentos por trecho.

A análise de referência foi realizada por duas médicas veterinárias que atuam na área de cardiologia com diferentes níveis de experiência: a primeira, com 23 anos de

atuação na área, e a segunda, com dois anos de prática clínica. Ambas avaliaram os 30 segmentos de forma independente, cegas à análise do ChatGPT, não havendo discordância entre as análises, estabelecendo-se assim o padrão-ouro definitivo.

O ChatGPT foi utilizado como ferramenta de apoio para a interpretação inicial dos mesmos segmentos, sem acesso ao padrão-ouro e sem treinamento prévio com os dados do estudo. Para cada trecho, o sistema forneceu a classificação do ritmo predominante, a identificação de eventos arrítmicos e, quando aplicável, a quantificação do número de eventos detectados. O desfecho primário do estudo foi a concordância entre ChatGPT e o padrão-ouro para a classificação do ritmo e para a detecção de eventos arrítmicos, incluindo SVPCs, VPCs, BAVs, pausas e alterações de frequência cardíaca como taquicardia sinusal e bradicardia sinusal. Como desfechos secundários, foram avaliadas a acurácia global na identificação de arritmias e a concordância quanto ao número de eventos detectados.

Para a análise estatística, foram utilizados diferentes métodos conforme o tipo de variável. A concordância categórica entre ChatGPT e padrão-ouro foi avaliada por meio do κ de Cohen. Para o desempenho diagnóstico, foram calculadas sensibilidade, especificidade, valor preditivo positivo (PPV), valor preditivo negativo (NPV) e F1-score. A análise de contagem de eventos foi realizada pelo método de Bland-Altman, com construção de limites de concordância entre o número de eventos identificados por cada abordagem. Todos os intervalos de confiança (IC95%) foram obtidos por bootstrap com 1.000 reamostragens, considerando a clusterização por animal. As análises foram realizadas no software GraphPad Prism®.

RESULTADOS, PROGRESSOS E REFLEXÕES

Entre os exames avaliados, a qualidade dos traçados foi considerada excelente. No padrão-ouro, o ritmo sinusal normal foi identificado em 40% dos segmentos, a arritmia sinusal em 30%, a taquicardia sinusal em 16,7%, a bradicardia sinusal em 6,7% e outros ritmos em 6,7%. Quanto à presença de eventos arrítmicos, foram detectadas SVPCs em 40% dos traçados, VPCs em 16,7%, bloqueios atrioventriculares Mobitz I/II em 10% e pausas sinusais ≥ 2 segundos em 6,7%. Não foram identificados eventos arrítmicos em 26,7% dos segmentos.

Quando comparado ao padrão-ouro, o ChatGPT apresentou sensibilidade de 80,5% e especificidade de 88,2% para a classificação do ritmo predominante, com acurácia global de 83,3% e κ de Cohen de 0,77, indicando concordância substancial. Para a detecção de eventos arrítmicos, a sensibilidade foi de 73,3%, a especificidade de 84,6%, a acurácia de 78,7% e o κ de Cohen de 0,71, também com concordância substancial.

Apesar do bom desempenho para eventos frequentes (SVPCs e VPCs), observou-se baixa sensibilidade para eventos críticos: 33% para bloqueio atrioventricular Mobitz I/II e 50% para pausas sinusais ≥ 2 s. Houve ainda cinco falso-positivos (três VPCs inexistentes e duas pausas não confirmadas). Assim, o estudo reforça a necessidade de supervisão manual por profissional da área e o uso da IA apenas como ferramenta auxiliar em cardiologia veterinária, sem substituir a avaliação de um médico veterinário, reforçando seu uso consciente.

CONSIDERAÇÕES FINAIS E PROJEÇÕES

Este estudo-piloto demonstrou que o ChatGPT apresentou bom desempenho na classificação do ritmo cardíaco e na identificação de arritmias comuns, mas com baixa sensibilidade para bloqueios atrioventriculares Mobitz I/II e pausas sinusais prolongadas, evidenciando limitações como único método diagnóstico. Embora a IA mostre potencial como ferramenta complementar na análise de Holter, os resultados reforçam a necessidade de supervisão por especialistas, especialmente em eventos raros e clinicamente relevantes. Estudos futuros com amostras maiores e treinamento específico de modelos de IA para sinais eletrocardiográficos são necessários para validar e aprimorar seu uso clínico.

Financiamento: Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio e incentivo à formação e qualificação profissional.

REFERENCIAS

BARRETTO, F.L. *et al.* Eletrocardiografia contínua (Holter) em cães saudáveis submetidos a diferentes exercícios físicos. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, [S.L.], v. 65, n. 6, p. 1625-1634, dez. 2013. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/s0102-09352013000600006>.

GÜNAY, Serkan *et al.* The accuracy of Gemini, GPT-4, and GPT-4o in ECG analysis: a comparison with cardiologists and emergency medicine specialists. **The American Journal Of Emergency Medicine**, [S.L.], v. 84, p. 68-73, out. 2024. Elsevier BV. <http://dx.doi.org/10.1016/j.ajem.2024.07.043>.

LIMA, Itamara Coelho *et al.* Avanços e Limitações da Inteligência Artificial Aplicada à Medicina Veterinária: diagnóstico e monitoramento animal. **Revista Contemporânea**, [S.L.], v. 5, n. 6, p. 1-23, 5 jun. 2025. Brazilian Journals. <http://dx.doi.org/10.56083/rev5n6-019>.

MAVROPOULOU, A.; OLIVEIRA, P.; WILLIS, R.. Holter monitoring in dogs: 24 h vs. 48 h. **The Veterinary Journal**, [S.L.], v. 272, p. 105628, jun. 2021. Elsevier BV. <http://dx.doi.org/10.1016/j.tvjl.2021.105628>.

MUZAMMIL, Muhammad Ali *et al.* Artificial intelligence-enhanced electrocardiography for accurate diagnosis and management of cardiovascular diseases. **Journal Of Electrocardiology**, [S.L.], v. 83, p. 30-40, mar. 2024. Elsevier BV. <http://dx.doi.org/10.1016/j.jelectrocard.2024.01.006>.

OSE, Benjamin *et al.* Artificial Intelligence Interpretation of the Electrocardiogram: a state-of-the-art review. **Current Cardiology Reports**, [S.L.], v. 26, n. 6, p. 561-580, 16 maio 2024. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s11886-024-02062-1>