



ESCOLA BAHIANA DE MEDICINA E SAÚDE PÚBLICA
CURSO DE MEDICINA

ANTONIO SANDRO SILVA ROCHA

**INTELIGÊNCIA ARTIFICIAL E MACHINE LEARNING NA PREDIÇÃO DE
PROGNÓSTICO EM PACIENTES COM INSUFICIÊNCIA CARDÍACA: UMA
REVISÃO SISTEMÁTICA**

TRABALHO DE CONCLUSÃO DE CURSO

SALVADOR - BA

2024

ANTONIO SANDRO SILVA ROCHA

**INTELIGÊNCIA ARTIFICIAL E MACHINE LEARNING NA PREDIÇÃO DE
PROGNÓSTICO EM PACIENTES COM INSUFICIÊNCIA CARDÍACA: UMA
REVISÃO SISTEMÁTICA**

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Medicina da Escola Bahiana de Medicina e Saúde Pública como requisito parcial para aprovação no 4º ano do curso de Medicina.

Orientador(a): Dr. Luiz Eduardo Fonteles Ritt

Coorientador(a): Dr. Gustavo Freitas Feitosa

SALVADOR

2024

Dedico este trabalho ao meu tio Rogério, que foi um segundo pai para mim, e que infelizmente não poderei compartilhar esse momento, mas ele estará sempre presente no meu coração.

AGRADECIMENTOS

Agradeço imensamente a Dra. Maria de Lourdes Lima pela paciência em ser minha tutora ao longo desses contratempos.

Ao Dr. Luiz Eduardo Fonteles Ritt por me aceitar como orientando e me dar oportunidades para além desse trabalho.

Ao Dr. Gustavo Freitas Feitosa por me ajudar e ficar à disposição durante todo esse trabalho.

Aos meus pais, Antonio Miranda Rocha e Sandra Silva Santos, por me apoiarem desde o início e por acreditarem na minha capacidade.

A minha namorada, Adryelle Thayne Araújo Linhares, por todo apoio e ajuda ao longo desse trabalho.

RESUMO

Introdução: A insuficiência cardíaca caracteriza-se por ser uma das principais causas de morbimortalidade no mundo. Nesse contexto, o cenário atual urge a necessidade de avaliar o prognóstico dos pacientes, e para isso são utilizados escores clássicos que possuem diversas limitações. Com o aumento do desenvolvimento da inteligência artificial e machine learning, espera-se usar essas ferramentas para superar essas limitações. **Objetivo:** Avaliar a performance de modelos de ML e IA na predição de prognóstico em pacientes com insuficiência cardíaca. **Metodologia:** Trata-se de revisão sistemática que buscou artigos nos principais banco de dados (MEDLINE, PubMed, Scielo e LILACS). Foram incluídos artigos que avaliam a predição da mortalidade ou readmissão em pacientes com insuficiência cardíaca através de modelos que utilizavam machine learning, deep learning ou inteligência artificial. Não houve delimitação de período dos estudos, bem como foram excluídos artigos que tratavam da previsibilidade diagnóstica. **Resultados:** Os resultados revelaram que há um crescente uso de algoritmos de ML para a análise de dados clínicos com resultados promissores na predição de desfechos de pacientes com IC, sobretudo para mortalidade (AUC média = 0.73 e IQR: 0.628-0.852), que desempenhou melhor quando comparado a predição de readmissão (AUC média = 0.64 e IQR: 0.595-0.707). Contudo, analisando de forma ampla, a performance se equiparou aos escores clássicos, com ressalva para alguns modelos de ML. **Conclusão:** apesar das dificuldades encontradas além de uma boa acurácia, a aplicação desses algoritmos fornece insights valiosos para melhorar a predição de prognóstico em pacientes com IC, entretanto, a utilização desses algoritmos deve ser avaliada com cuidado em virtude das inúmeras abordagens e aplicações diferentes apresentadas por cada autor.

Palavras-chave: Insuficiência cardíaca. Predição de prognóstico. Inteligência artificial. Machine learning. Deep learning.

ABSTRACT

Introduction: Heart failure is characterized as one of the leading causes of morbidity and mortality worldwide. In this context, the current scenario urgently calls for evaluating patient prognosis, and for this purpose, classical scores with several limitations are commonly used. With the increasing development of artificial intelligence (AI) and machine learning (ML), there is an expectation that these tools can overcome these limitations. **Objective:** To evaluate the performance of ML and AI models in predicting prognosis in patients with heart failure. **Methodology:** This systematic review searched articles in major databases (MEDLINE, PubMed, Scielo, and LILACS). Included articles assessed mortality or readmission prediction in heart failure patients using models based on machine learning, deep learning, or artificial intelligence. There were no restrictions on the study period, and articles related to diagnostic predictability were excluded. **Results:** The results revealed a growing use of ML algorithms for analyzing clinical data, showing promising outcomes in predicting patient outcomes, particularly mortality (average AUC = 0.73, IQR: 0.628-0.852), which outperformed readmission prediction (average AUC = 0.64, IQR: 0.595-0.707). However, when analyzed broadly, the performance was comparable to classical scores, with some ML models showing exceptions. **Conclusion:** Despite encountered difficulties, the application of these algorithms provides valuable insights for improving prognosis prediction in heart failure patients. However, the use of these algorithms should be carefully evaluated due to the numerous different approaches and applications presented by each author.

Keywords: Heart failure. Prognosis prediction. Artificial intelligence. Machine learning. Deep learning.

LISTA DE ABREVIATURAS E SIGLAS

ML	Machine Learning
IA	Inteligência Artificial
HF	Heart Failure
LR	Logistic Regression
RF	Random Forest
XGBoost	Extreme Gradient Boosting
GBM	Gradient Boosting Machines
SVM	Support Vector Machine
LASSO	Least Absolute Shrinkage and Selection Operator
MVE-EK	Multi-view Ensemble Learning with Empirical Kernel
FRDLS	Feature Rearrangement based Deep Learning System
MARKER	Machine learning Assessment of Risk and Early mortality

SUMÁRIO

1	INTRODUÇÃO	7
2	OBJETIVOS	9
2.1	Geral	9
2.2	Específicos	9
3	REVISÃO DA LITERATURA	10
4	MÉTODOS	13
4.1	Desenho do Estudo	13
4.2	Período	13
4.3	Critérios de Elegibilidade	13
4.4	Fontes de Informação	13
4.5	Estratégia de Busca	13
4.6	Seleção e Processo de Coleta dos Estudos	14
4.7	Avaliação da Qualidade dos Estudos	14
4.8	Extração dos Dados	15
4.9	Análise Estatística	15
4.10	Aspectos Éticos	15
5	RESULTADOS	16
5.1	Identificação e Seleção	16
5.2	Características dos Estudos Incluídos	17
5.3	Qualidade dos Estudos	21
5.4	Principais Achados dos Estudos	21
5.5	Comparativo com Escores Clássicos	28
6	DISCUSSÃO	30
7	CONCLUSÃO	32
	REFERÊNCIAS	34

1 INTRODUÇÃO

A insuficiência cardíaca (IC) é uma condição clínica bastante complexa, devido aos diversos fatores de riscos envolvidos, bem como pela natureza dessa síndrome: incapacidade do coração de bombear o sangue adequadamente para suprir a necessidade metabólica do indivíduo.¹ Além de ser uma condição bastante complexa, a insuficiência cardíaca caracteriza-se por ser uma das principais causas de morbimortalidade no mundo, principalmente no Brasil, onde a incidência de novos casos podem chegar a 240 mil por ano. Muitos fatores de riscos precedem a IC, destacando-se a idade como importante vetor resultante para esta síndrome. Em virtude disso, segundo dados do IBGE, ao longo de 10 anos a população idosa brasileira cresceu aproximadamente 40%, além disso, estima-se que até 2100 o índice de envelhecimento da população brasileira será o dobro da média mundial, evidenciando uma preocupação no aumento de casos de IC no Brasil.²⁻⁴

Em paralelo com o aumento da prevalência de IC, o âmbito tecnológico vem crescendo bastante, destacando-se a aplicação de inteligência artificial e *machine learning* em diversas áreas como ferramentas facilitadoras e de maior precisão quando comparada a análise de dados feita de forma manual, uma vez que é possível trabalhar de forma mais robusta e automatizada. Em virtude disso, essas ferramentas têm um enorme potencial no campo da medicina, principalmente quando se trata de predições, seja diagnóstica ou prognóstica.⁵

Todavia, atualmente, para suprir essa necessidade de predizer o prognóstico dos pacientes com IC, diversos escores com diferentes abordagens paramétricas são utilizados, e dentre os principais estão: *Heart Failure Survival Score*, *Seattle Heart Failure Model*, *Metabolic Exercise Cardiac Kidney Index*, *Meta-analysis Global Group in Chronic Heart Failure*. Nesse contexto, alguns desses escores possuem melhor acurácia para hospitalização e outros para mortalidade, sendo necessário muitas vezes cruzar informações entre eles para melhor avaliação do paciente.⁶⁻⁸

Em virtude disso, é importante um estudo que analise cuidadosamente as incertezas sobre a capacidade das inteligências artificiais, bem como de *machine learning*, para realizar uma análise mais abrangente, processando uma grande quantidade de dados, cruzando parâmetros e fazendo novas relações preditivas, principalmente sobre

hospitalizações e mortalidade, que às vezes não são alcançadas ou são limitadas aos escores do estado da arte. De igual modo, esses recursos possibilitam também a promoção de informações para otimização dos protocolos e um melhor acompanhamento dos pacientes com essa comorbidade.⁹

2 OBJETIVOS

2.1 Geral

Avaliar a performance de algoritmos de inteligência artificial e *machine learning* na predição de prognósticos em pacientes com insuficiência cardíaca.

2.2 Específicos

1. Comparar a capacidade dos escores clássicos com a inteligência artificial na predição de desfechos como mortalidade e hospitalização em pacientes com insuficiência cardíaca;
2. Analisar a acurácia e a ferramenta utilizada pelos modelos de inteligência artificial na perspectiva de predição.

3 REVISÃO DA LITERATURA

A proeminência da inteligência artificial na sociedade contemporânea se apresentou e trouxe novamente as faces do mundo moderno como uma experiência mais elaborada, firme e renovada das resoluções e perspectivas fracassadas durante os anos setenta para a medicina.

Ela nasce juntamente com o movimento renovador do papel da tecnologia e comunicação para a sociedade e do próprio computador em si. Castells estende e observa que a tecnologia de informação sobre o papel das redes no mundo e a incorporação aos domínios da sociedade humana num todo. A perspectiva promissora durante os anos de 1959 com a IA para a medicina eram muitas, William B. Schwartz apresentou que a tecnologia e a ciência da computação em trinta anos iriam substituir as funções intelectuais dos médicos, o que não acabou se provando concreto devidos aos percalços de desenvolvimento material, sendo estes entraves as condições de velocidade de rede e processamento de informações, bem como quanto ao papel de simular a função dos médicos nos atendimentos clínicos. Dessa forma, culminou no questionamento de Schwartz para a tecnologia e as tamanhas expectativas dela auxiliar o papel da medicina e do atendimento clínico.⁹

Com as mudanças constantes na evolução tecnológica depois dos anos 2000, a IA e a tecnologia se tornaram mais compacta e mais poderosas, com efeitos mais imediatos e capacidades de coletar numerosas quantidades de dados comparado ao século XX. Os instrumentos para mensuração das condições clínicas se tornaram uma via importante para dentro dos setores clínicos, e o seu aprendizado se passa a ser importante juntamente quando unido a prontuários médicos possibilitando identificação de padrões e condições não tão fáceis de observação.¹⁰ Em particular para as doenças cardiovasculares que provocam 17 milhões de mortes no mundo e dada a importância de um órgão vital como o coração, a necessidade da previsão se tornou uma prioridade na prática clínica.

Com o nível de mortalidade que as doenças cardiovasculares causam, com taxas estimadas em 31,5% de mortes da população mundial no ano de 2015, o uso de noções de big data, programação e dados que possibilitam analisar conjuntos numerosos tornou-se prioridade atualmente, uma vez que até o momento a tecnologia vigente não é capaz de elaborar previsões em variáveis tão precisas, além de

possuírem baixa possibilidade de interpretação. Alguns autores ainda apresentam que a pontuação para as problemáticas em relação aos diagnósticos é a dificuldade, assim como a falta da reprodutibilidade que acaba sendo um dos grandes determinantes para ter diagnósticos e prognósticos com fineza exatidão.¹¹

A inteligência artificial com os níveis alcançados em 2023, com a possibilidade de simplificar e acelerar os ensaios clínicos através de sistematizações que constroem modulações mais eficientes e com a correspondência dos participantes dos estudos, a IA e suas inúmeras possibilidades de trabalhar com diagnósticos e prognóstico, pode realizar múltiplas tarefas devido seu processamento muito mais avançado de machine learning”, o que possibilita a IA construir pacientes sintéticos e analisar as possíveis reações adversas em subgrupos e populações focais.¹²

A compreensão do uso da IA na medicina se dá por determinados fatores que hoje são uma questão vigente no poderio medicinal, nos fazendo pensar sobre a possibilidade da inclusão da inteligência artificial no meio clínico. Nos promove a pensar que a IA apesar da pouca adesão de médicos, ela pode servir como um motor para acelerar diagnósticos, verificar melhores versões de medicamentos e entender o percurso clínico do paciente numa velocidade e simplismo que demandaria uma equipe muito grande para cada paciente, ao qual logisticamente não é possível.

Os dados aprendidos em machine learning pode auxiliar diagnósticos mais tarde, entretanto há suas problemáticas, os primeiros pontos são na construção da coleta, é necessária uma grande quantidade de dados. Então sua precisão é diretamente posicionada com base na quantidade de dados coletados, e o segundo ponto é a disposição desses dados na interpretação. A capacidade da machine learning é a possibilidade de visualizar algo que o ser humano não consegue observar devido o número excessivo de dados, ou seja, a visualização dos resultados é um papel inquestionável do que tange ao machine learning, mas a incapacidade da realização de cruzamentos mais complexos e principalmente no que se refere a situações mais abstratas.

É importante salientar, que a incapacidade de compreender a lógica por trás de um modelo preditivo pode causar problemas para o médico interpretar e utilizar essas previsões em sua prática. Um exemplo em que uma machine learning desempenhou resultados excelentes, todavia, careceu de lógica clínica: este estudo investigou a

predição de mortalidade em pacientes com pneumonia, para que tratassem os de baixo risco ambulatorialmente ou de alto risco intra-hospitalar. O modelo previu que pacientes com asma e pneumonia tinham menor risco de morte quando comparado aos pacientes com apenas pneumonia. Logicamente, é um dado contraintuitivo, mais isso ocorreu porque os pacientes com asma foram internados como medida de precaução, resultando em menor mortalidade, enquanto os pacientes não asmáticos e que não foram internados, apresentaram maior mortalidade. Isto é, um algoritmo, mesmo que bem estruturado e com boa precisão, não tem esse tipo de concepção ou previsibilidade lógica que somente os humanos conseguem fazer, nem a capacidade de entender o significado de cada dado alimentado.¹³

Logo, destaca-se a necessidade de uma abordagem cautelosa para implementar ML na prática clínica, pois deve-se garantir que o contexto seja analisado para que os resultados sejam interpretados corretamente. Apesar dos desafios enfrentados, é importante ressaltar os benefícios da IA em questão na medicina, visto que a aplicação de algoritmos de ML para as predições em readmissões de pacientes com insuficiência cardíaca tem sido amplamente explorada nos últimos anos.¹⁴

Além de auxiliar na predição da readmissão, possibilita que medidas preventivas sejam tomadas, com o intuito de melhorar desfechos clínicos e podendo levar a uma redução significativa nas taxas de readmissão e custos associados ao tratamento.¹⁵ Nesse sentido, as pesquisas exploratórias devem focar também na interpretação e aplicabilidade clínica, além da precisão, uma vez que a aplicação de IA e ML na predição de prognóstico em pacientes com insuficiência cardíaca oferece diferentes possibilidades e diferentes abordagens em um vasto campo multidisciplinar, a fim de explorar e validar tais ferramentas para a melhoria de condições médicas.

4 MÉTODOS

4.1 Desenho do Estudo

Trata-se de uma revisão sistemática da literatura baseada nas recomendações do guia *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA 2020 STATEMENT).

4.2 Período

Não foi delimitado período específico, sendo elegíveis estudos realizados em qualquer ano.

4.3 Critérios de Elegibilidade

Os trabalhos considerados elegíveis para este estudo foram ensaios clínicos randomizados e estudos de coorte publicados completamente em inglês, português ou espanhol, realizados em qualquer ano, e que envolve população com insuficiência cardíaca acompanhadas por uma inteligência artificial ou *machine learning* na perspectiva da predição de prognóstico.

Foram excluídos estudos que tratam da predição diagnóstica, classificação de risco ou que a temática não seja relevante para este estudo.

4.4 Fontes de Informação

Este estudo coletou amostras por meio de uma busca virtual nas bases de dados Embase, Medlars Online (MEDLINE), Public Medline (PUBMED), Literatura Latino-Americanae do Caribe em Ciências da Saúde (LILACS) e Scientific Electronic Library Online (SciELO). Ademais, a fim de ampliar as buscas, foram analisadas as referências dos artigos selecionados na tentativa de encontrar outros estudos que preencham os critérios de inclusão.

4.5 Estratégia de Busca

As buscas ocorridas foram realizadas no modo avançado pelos títulos em cada uma das fontes citadas anteriormente. Em um primeiro momento foram selecionados

os principais descritores relacionados ao tema a partir do *MeSH Heading* e DeCS, provenientes da *National Library of Medicine* e Biblioteca Virtual em Saúde respectivamente. Os descritores encontrados e selecionados foram: *Heart Failure*, *Adult congenital disease*, *Artificial Intelligence*, *Machine Learning*, *Deep Learning*.

Em seguida, com intuito de filtrar as informações coletadas, foi utilizado operadores booleanos OR e AND nos campos de busca avançados, começando por descritores mais amplos até os mais específicos. A seguinte estrutura dos descritores foi utilizada na busca:

("heart failure" OR "Adult congenital heart disease") AND ("artificial intelligence" OR "machine learning" OR "machine intelligence" OR "artificial intelligence")

A identificação dos estudos foi realizada com base na estratégia PECO, acrônimo para *Population* (P), *Exposition* (E), *Comparison* (C) e *Outcomes* (O):

P: Pacientes com Insuficiência Cardíaca

E: Algoritmos e escores baseados em Inteligência Artificial e *Machine Learning*

C: Escores clássicos considerados estados da arte

O: Hospitalização e Mortalidade

4.6 Seleção e Processo de Coleta dos Estudos

Os estudos foram selecionados por um revisor de forma independente, utilizando o software "Rayyan", compreendendo os critérios de inclusão e a delimitação temática, bem como por triagem: remoção dos estudos duplicados; análise dos títulos e resumo a fim de excluir os artigos não elegíveis; leitura na íntegra dos artigos escolhidos, com aprofundamento na aplicação dos critérios de elegibilidade, para filtrar os estudos com maior relevância e qualidade que dialogam com objetivo e temática deste estudo. Além disso, para auxiliar na seleção dos artigos, foi utilizado um filtro baseado na PECO, disponibilizado pelo software "Rayyan".

4.7 Avaliação da Qualidade dos Estudos

Com propósito de garantir a qualidade da amostra selecionada para esta revisão sistemática, utilizamos a ferramenta *New Castle Ottawa* (NOS) para avaliar os estudos observacionais.

4.8 Extração dos Dados

A extração dos dados também foi realizada por um pesquisador independente. O registro dos dados foi norteado com base em um formulário pré-definido. Os parâmetros coletados inicialmente serão relativos ao artigo: título, resumo, tipo de estudo, ano e país de publicação e referências. A partir disso, o enfoque da extração foi voltado para perfil clínico dos pacientes, bem como para os recursos utilizados como inteligência artificial ou *manchine learning*, e principalmente dados no que tange os desfechos primários.

4.9 Análise Estatística

Em virtude da temática, uma tabela-síntese foi montada para evidenciar os principais resultados da acurácia na predição de prognóstico das inteligências artificiais e machine learning, valendo-se de uma análise através da área da curva ROC (AUC) para cada desfecho analisado, bem como pelos valores de Recall (sensibilidade), acurácia e F-score quando fornecidos pelo autor.

4.10 Aspectos Éticos

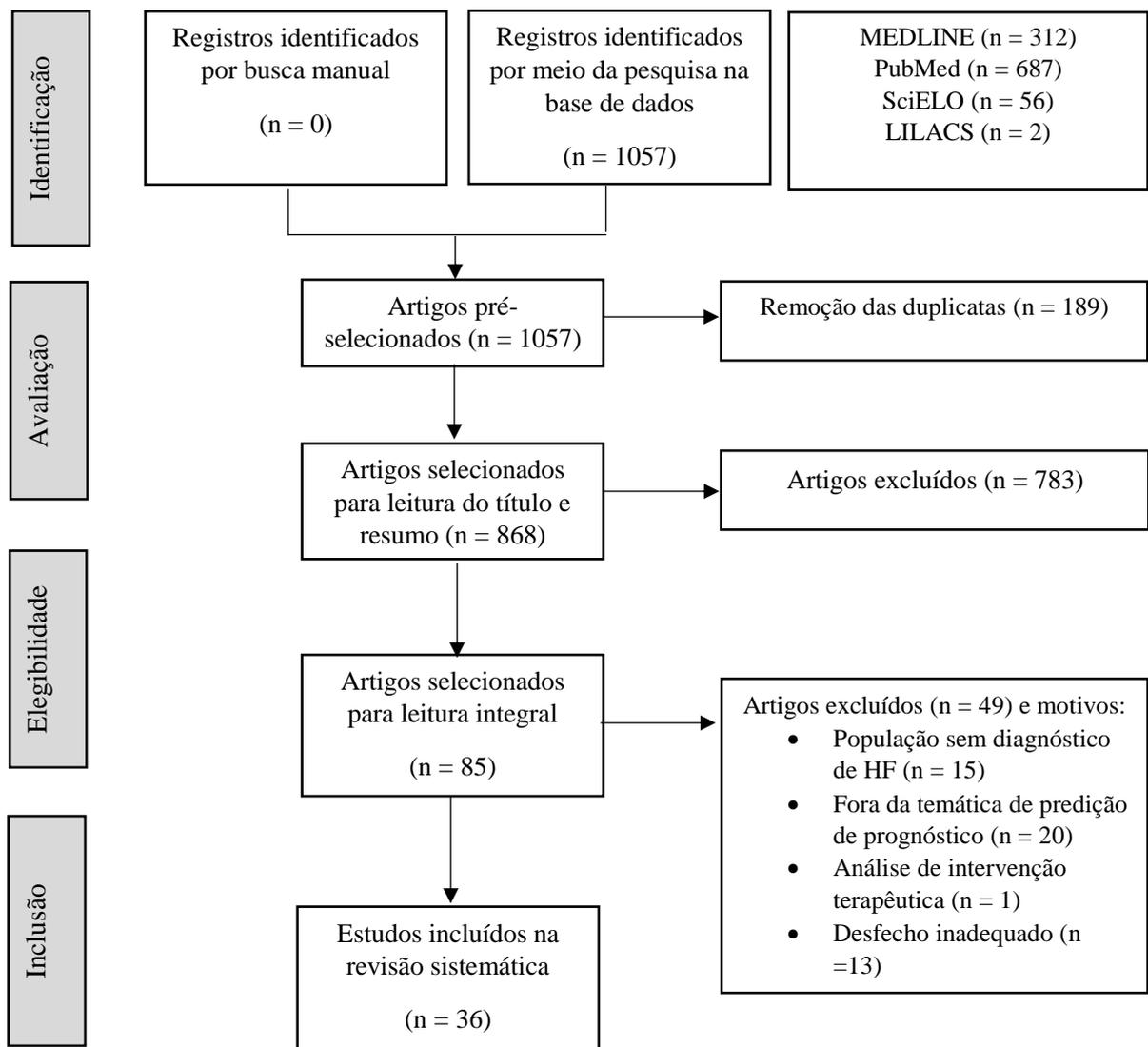
Caracterizado por uma revisão da literatura, este estudo não foi submetido ao Comitê de Ética em Pesquisa de acordo com a Resolução 466/12 do Conselho Nacional de Saúde (CNS), que garante a isenção do termo esclarecido de consentimento e submissão ao CEP. Todavia, a condução desse estudo respeitará todos os preceitos éticos, sendo conduzido de forma responsável, com informações clara e precisas, zelando pela transparência e legitimidade.

5 RESULTADOS

5.1 Identificação e Seleção

Através da pesquisa nas bases de dados, foram identificados 1057 registros, dos quais 189 eram duplicatas e permanecendo 868 registros. A partir do processo de triagem por meio do título e resumo, foram selecionados 85 estudos para leitura na íntegra, dos quais 49 foram excluídos com base nos critérios de elegibilidade. Por fim, restaram apenas 36 estudos que foram incluídos para a análise qualitativa, como é possível observar na esquematização da Figura 1.

Figura 1 - Fluxograma da seleção dos artigos (de acordo com o PRISMA)



Fonte: próprio autor

5.2 Características dos Estudos Incluídos

Dos 36 estudos encontrados, todos tem como desenho de estudo coorte prospectiva, boa parte dos estudos analisaram a predição de mortalidade, apenas 17 deles analisaram somente a mortalidade, 13 somente a readmissão, e 6 analisaram os dois desfechos de readmissão e mortalidade, destacando 2 estudos que analisaram de forma combinada na machine learning, isto é, os desfechos não foram separados e a máquina tentava predizer o acontecimento de um desfecho ou outro simultaneamente. (Tabela 1).

Tabela 1 - Características gerais dos estudos selecionados

Autores	Período dos dados	Fonte dos dados	Nº de pacientes	Desfecho avaliado
Frizzell J. D. (2016) ¹⁶	2005-2010	Registros	56.477	Readmissão em 30 dias
Wang Z. (2020) ¹⁷	2009-2016	PE	10.203	Mortalidade em 1 ano
Greenberg B. (2021) ¹⁸	-	PE	4064	Mortalidade*
Mahajan S. M. (2018) ¹⁹	2015-2015	PE	1778	Readmissão em 30 dias
Tian P. (2023) ²⁰	-	Registros	424	Mortalidade*
Tohyama T. (2021) ²¹	2013-2013	Claims dataset	10.175	Mortalidade em 1 ano
Austin D.E. (2022) ²²	2004-2007	Registros	12.608	Mortalidade 30 dias
Adler E. D. (2019) ²³	-	PE	5822	Mortalidade 1 ano
Ortiz J. (1995) ²⁴	-	Banco de dados e Ecocardiograma	95	Mortalidade em 1 ano
Negassa A. (2021) ²⁵	2001-2010	PE	7516	Mortalidade em 30 dias
Sharma V. (2022) ²⁶	2012-2019	Claims dataset	9845	Readmissão em 30 dias
Mahajan SM (2016) ²⁷	2011-2015	PE	36.245	Readmissão em 30 dias
Gutman R. (2022) ²⁸	2004-2015	PE	10.868	Mortalidade em 1 ano
Li D. (2023) ²⁹	2001-2012	Banco de dados público	10.311	Mortalidade em 1 ano
Jhangiri S. (2024) ³⁰	2020-2020	Banco de dados de readmissões	489.482	Readmissão em 30 dias
Allam A. (2019) ³¹	-	Claims dataset	272.778	Readmissão em 30 dias
Beecy N. A. (2017) ¹⁰	-	PE	3774	Readmissão ou mortalidade em 30 dias
Ju C. (2021) ³²	2013-2017	PE	8893	Mortalidade em 30 dias
Ru B. (2023) ³³	2013-2017	PE	30.687	Readmissão em 30 dias

Soliman A. (2023) ³⁴	2017-2019	Registros e PE	15.612	Readmissão em 100 dias
Chen S. (2023) ³⁵	2016-2019	PE	2002	Readmissão em 6 meses
Ashfaq A. (2019) ³⁶	2012-2016	PE	7.500	Readmissão em 30 dias
Angraal S. (2019) ³⁷	-	Ensaio clínico	1.767	Readmissão ou mortalidade em 3 anos
Xu C. (2023) ³⁸	2017-2019	Registros médicos	622	Mortalidade em 3 anos
König S (2021) ³⁹	2016-2018	Banco de dados multicêntrico	59.125	Mortalidade*
Ketabi M. (2024) ⁴⁰	2015-2020	Registros	2488	Mortalidade em 1 ano
Kwon J. (2019) ¹⁵	2016-2017	Registro multicêntrico	8094	Mortalidade em 1 ano
Shameer al (2017) ⁴¹	-	PE	1068	Readmissão em 30 dias
Lv H. (2021) ⁴²	2010-2018	PE	13.602	Readmissão ou mortalidade em 1 ano
Sarijaloo F. (2021) ⁴³	2011-2019	PE	3189	Readmissão ou mortalidade em 90 dias
van der Galiën (2021) ⁴⁴	2012-2014	Claims dataset	25.776	Readmissão ou mortalidade em 1 ano
Awan SE (2019) ⁴⁵	2003-2008	Claims dataset	10.757	Readmissão ou mortalidade em 30 dias
Wang Z. (2021) ⁴⁶	2009-2016	PE	4682	Mortalidade em 1 ano
Jing L. (2020) ⁴⁷	2001-2019	PE	26.971	Mortalidade em 1 ano
Najafi-Vosough R. ⁴⁸ (2021)	2015-2019	Registros médicos	1856	Readmissão*
Sabouri M. (2023) ⁴⁹	-	Registros	737	Readmissão em 30 dias

Fonte: próprio autor

Legenda: PE – Prontuário eletrônico

* Não houve um período específico

Ainda podemos notar que a maioria optou por utilizar prontuários eletrônicos como fonte de dados para alimentar a machine learning e estipular as variáveis necessárias para predição. Dos 36 estudos realizados, 17 utilizaram exclusivamente prontuários eletrônicos, 5 usaram somente registros do hospital e 5 *claims datasets* (conjunto de dados que pode conter dados de prontuário como também registros administrativos dos hospitais). Os demais estudos se distribuíram entre registros médicos, registros multicêntricos, bancos de dados, ensaios clínicos e ecocardiograma. Vale destacar a quantidade de pacientes em cada estudo, que se mostrou bastante considerável, com mediana 7805 (IQR: 2245-10812.5), como pode ser visualizado na tabela 2.

Tabela 2 - Características demográficas dos pacientes com Insuficiência Cardíaca

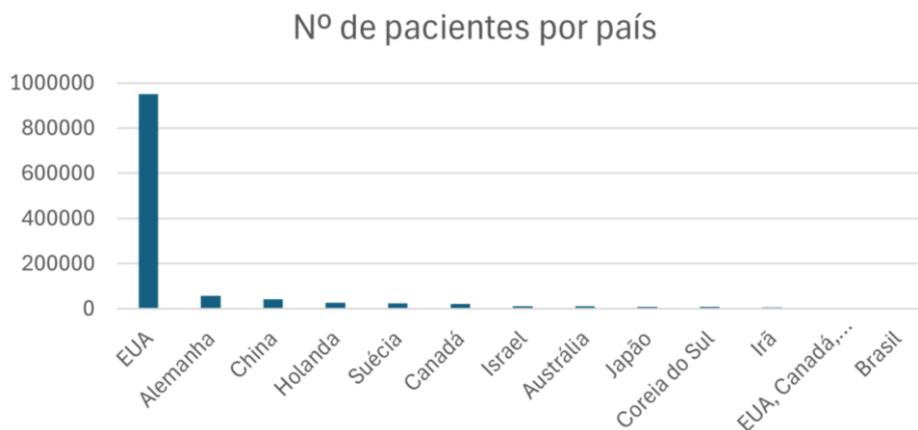
Autores	Região do Estudo	Nº de pacientes	Idade	% Gênero masculino
Frizzell J. D. (2016) ¹⁶	EUA	56.477	65	-
Wang Z. (2020) ¹⁷	China	10.203	-	-
Greenberg B. (2021) ¹⁸	EUA	4064	59 ± 13	42
Mahajan S. M. (2018) ¹⁹	EUA	1778	72.3 ± 12.1	97.6
Tian P. (2023) ²⁰	China	424	59.34 ± 14.25	71.2
Tohyama T. (2021) ²¹	Japão	10.175	80 [70-86]	54.7
Austin D.E. (2022) ²²	Canadá	12.608	-	-
Adler E. D. (2019) ²³	EUA	5822	60.3	36.7
Ortiz J. (1995) ²⁴	Brasil	95	51.25 ± 12.27	60
Negassa A. (2021) ²⁵	EUA	7516	71.8	45.04
Sharma V. (2022) ²⁶	Canadá	9845	71.5	56
Mahajan SM (2016) ²⁷	EUA	36.245	-	-
Gutman R. (2022) ²⁸	Israel	10.868	74.4 ± 12.2	49.1
Li D. (2023) ²⁹	EUA	10.311	-	-
Jahangiri S. (2024) ³⁰	EUA	489.482	71	-
Allam A. (2019) ³¹	EUA	272.778	72.89 ± 14	51
Beecy N. A. (2017) ¹⁰	EUA	3774	-	-
Ju C. (2021) ³²	China	8893	81	45.27
Ru B. (2023) ³³	EUA	30.687	70.1 ± 13	59.33
Soliman A. (2023) ³⁴	Suécia	15.612	79.1 ± 10.4	55.77
Chen S. (2023) ³⁵	China	2002	-	48
Ashfaq A. (2019) ³⁶	Suécia	7.500	78.8	57
Angraal S. (2019) ³⁷	EUA, Canadá, Argentina e Brasil	1.767	72 [64-69]	50.1
Xu C. (2023) ³⁸	China	622	66.27 ± 12.25	64.2
König S (2021) ³⁹	Alemanha	59.125	75	48.1
Ketabi M. (2024) ⁴⁰	Irã	2488	65.68 ± 13.22	63.1
Kwon J. (2019) ¹⁵	Coreia do Sul	8094	68.2	52.7
Shameer al (2017) ⁴¹	EUA	1068	-	-
Lv H. (2021) ⁴²	China	13.602	74.5	52.5
Sarijaloo F. (2021) ⁴³	EUA	3189	67.8 ± 15	-
van der Galiën (2021) ⁴⁴	Holanda	25.776	74 (66– 80)	56
Awan SE (2019) ⁴⁵	Austrália	10.757	82.1	49
Wang Z. (2021) ⁴⁶	China	4682	-	-
Jing L. (2020) ⁴⁷	EUA	26.971	76 (67– 84)	53
Najafi-Vosough R. ⁴⁸ (2021)	Irã	1856	71.7	-
Sabouri M. (2023) ⁴⁹	Irã	737	55.85	71.8

Fonte: próprio autor

Observa-se que a idade média dos pacientes que entraram nos estudos reflete uma população idosa, com média de 70 anos, considerando um cálculo dos valores centrais das idades. Entretanto, esse cálculo tem suas devidas limitações, visto que nem todos os autores forneceram dados sobre a idade, bem como alguns não forneceram o intervalo e desvio padrão. Em relação ao gênero, nota-se maioria composta pelo gênero masculino, em média 55.16% dos pacientes (IQR: 51-64).

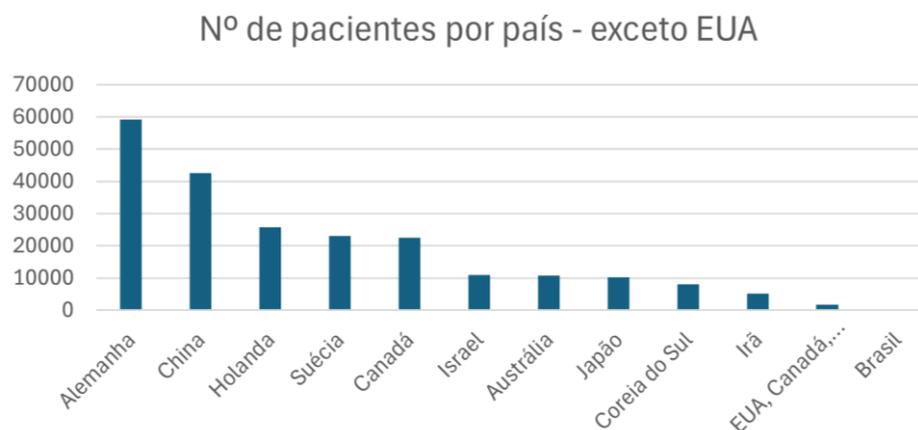
Salienta-se que muitos desses estudos ocorreram em países mais desenvolvidos como os Estados Unidos ($n = 14$), seguido da China ($n = 8$), Irã ($n = 3$), Suécia ($n = 2$) e Canadá ($n = 2$). Não somente isso como também há um número de pacientes discrepante entre estudos realizados nos Estados Unidos e nos demais países, como é possível observar na Figura 2.

Figura 2 – Número da população estudada por país



Fonte: próprio autor

Figura 3 – Número da população estudada por país, exceto EUA



Fonte: próprio autor

5.3 Qualidade dos Estudos

Em sua maioria, os estudos obtiveram uma boa pontuação na escala de Newcastle Ottawa, com ressalva para alguns artigos que não explicitaram se houve controle de fator de confusão, bem como quais fatores foram controlados (Tabela 3). Embora subtenda-se que na utilização de ML e IA costuma-se controlar fatores de confusão relacionados aos dados demográficos, há outros fatores muito importantes que precisam ser controlados assim como associar métodos diferentes para mitigar o viés.

Tabela 3 - Avaliação dos estudos com Newcastle Ottawa (NOS) ordenados pelo score

Autor	Seleção	Comparabilidade	Desfecho	Score
Frizzell J. D. (2016) ¹⁶	★★★★	★★	★★★	9
Wang Z. (2020) ¹⁷	★★★★	★★	★★★	9
Adler E. D. (2019) ²³	★★★★	★★	★★★	9
Negassa A. (2021) ²⁵	★★★★	★★	★★★	9
Sharma V. (2022) ²⁶	★★★★	★★	★★★	9
Gutman R. (2022) ²⁸	★★★★	★★	★★★	9
Jahangiri S. (2024) ³⁰	★★★★	★★	★★★	9
Beecy N. A. (2017) ¹⁰	★★★★	★★	★★★	9
Ju C. (2021) ³²	★★★★	★★	★★★	9
Soliman A. (2023) ³⁴	★★★★	★★	★★★	9
Chen S. (2023) ³⁵	★★★★	★★	★★★	9
Kwon J. (2019) ¹⁵	★★★★	★★	★★★	9
Shameer al (2017) ⁴¹	★★★★	★★	★★★	9
Lv H. (2021) ⁴²	★★★★	★★	★★★	9
Sarijaloo F. (2021) ⁴³	★★★★	★★	★★★	9
Wang Z. (2021) ⁴⁶	★★★★	★★	★★★	9
Najafi-Vosough R. (2021) ⁴⁸	★★★★	★★	★★★	9
Sabouri M. (2023) ⁴⁸	★★★★	★★	★★★	9
Allam A. (2019) ³¹	★★★★	★	★★★	8
Ru B. (2023) ³³	★★★★	★	★★★	8
Ashfaq A. (2019) ³⁶	★★★★	★	★★★	8
Angraal S. (2019) ³⁷	★★★★	★★	★★	8
Xu C. (2023) ³⁸	★★★★	★	★★★	8
König S (2021) ³⁹	★★★★	★	★★★	8
Ketabi M. (2024) ⁴⁰	★★★★	★	★★★	8
van der Galiën (2021) ⁴⁴	★★★★	★	★★★	8
Awan SE (2019) ⁴⁵	★★★★	★	★★★	8
Jing L. (2020) ⁴⁷	★★★★	★	★★★	8
Greenberg B. (2021) ¹⁸	★★★★	★	★★	7
Tian P. (2023) ²⁰	★★★★	★	★★	7
Tohyama T. (2021) ²¹	★★★★	-	★★★	7
Mahajan S. M. (2018) ¹⁹	★★★★	-	★★	6
Austin D.E. (2022) ²²	★★★★	-	★★	6
Mahajan SM (2016) ²⁷	★★★★	-	★★	6
Li D. (2023) ²⁹	★★★★	-	★★	6
Ortiz J. (1995) ²⁴	★★★	★	★	5

Fonte: próprio autor.

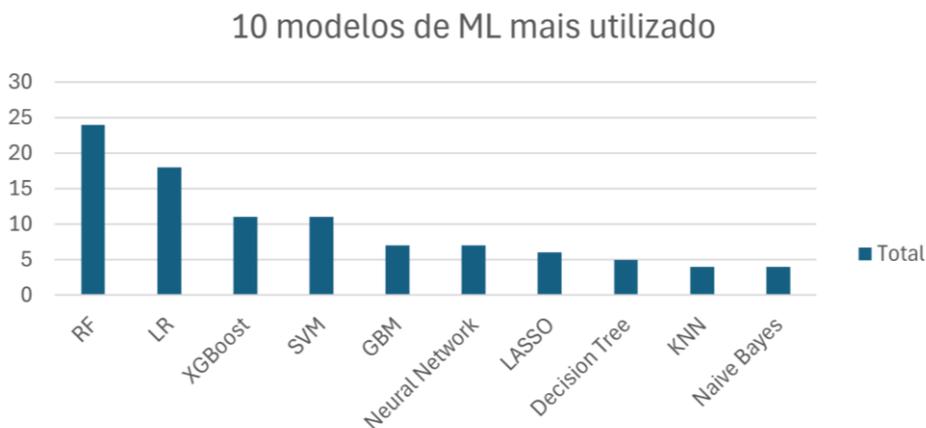
Apenas 5 artigos obtiveram um score mediano na escala de NOS, destacando-se para o estudo de Ortiz J. que obteve o score mais baixo.

5.4 Principais Achados dos Estudos

A maioria dos estudos avaliaram a predição com ML e DL, esse primeiro é um tipo de tecnologia considerado um subconjunto de uma inteligência artificial, em que desempenha um papel semelhante a um motor da IA que processa os dados e consegue aprimorá-los, enquanto o segundo é o aprofundamento do primeiro com redes neurais. Se tratando de uma temática muito nova, espera-se que os estudos comecem por uma machine learning com intuito de obter uma validação, e após isso, evoluírem para inteligência artificial.

Considerando que há diversas variações dos modelos principais de ML, 168 modelos foram usados para analisar a predição de prognóstico em pacientes com IC, destacando-se para os 10 principais (Figura 4).

Figura 4 – Principais modelos utilizados nos estudos



Fonte: próprio autor

Dentre esses mais utilizados, destaca-se bom resultado na predição de mortalidade para XGBoost (AUC média = 0.810) e GBM (AUC média = 0.800). Alguns modelos foram muito satisfatórios, como por exemplo MARKER-HF, MVE-K, FRDLS, LASSO e SVM. Contudo, são estudos isolados e por isso não se pode concluir sua total eficácia.

Tabela 4 - Principais achados dos estudos

Autor	Algoritmos	Sensibilidade	Acurácia	AUC (mortalidade)	AUC (readmissão)	F-score
Frizzell J. D. (2016) ¹⁶	TAN				0.62	
	GBM				0.61	
	LASSO				0.62	
	LR				0.62	
	RF				0.61	
Wang Z. (2020) ¹⁷	FRDLS	72.1	80.31	0.886		0.690
	CS-SVM	72.9	80.82	0.808		0.713
	MLP	56.5	74.56	0.881		0.641
	Perceptron	55.2	72.19	0.852		0.689
	LR	70.1	70.70	0.782		0.536
	RF	45.3	71.26	0.891		0.592
	XGBoost	66.2	79.52	0.903		0.713
	LightGBM	64.5	78.02	0.883		0.689
Greenberg B. (2021) ¹⁸	OS-SVM	67.9	78.36	0.814		0.686
	MARKER-HF			0.85 [0.83-0.88]		
Mahajan S. M. (2018) ¹⁹	Boosted Trees				0.71 [0.57-0.72]	
	Spike-and-Lab (g-net)				0.60 [0.59-0.61]	
Tian P. (2023) ²⁰	XGBoost	84.3	0.854	0.91 [0.88– 0.94]		0.767
	RF	83.5	0.821	0.90 [0.87– 0.93]		0.727
	SVM	77.7	0.781	0.80 [0.75– 0.85]		0.669
	Logistic lasso	66.8	73.3	0.77 [0.75–0.80]		
	SVM	69.7	69.6	0.76 [0.73–0.79]		
Tohyama T. (2021) ²¹	RF	66.5	72.4	0.76 [0.74–0.79]		
	Gradient boosting tree	61.0	75.0	0.76 [0.73–0.79]		
	ACD-VC	71.1	70.7	0.77 [0.75–0.80]		
	Neural Network	66.8	73.8	0.77 [0.75–0.80]		
	LASSO			0.74 [0.72-0.76]		
	Neural Network			0.68 [0.65-0.71]		
Austin D.E. (2022) ²²	RF			0.74 [0.72-0.76]		
	XGBoost			0.75 [0.74-0.77]		
	MARKER-HF			0.88		
Adler E. D. (2019) ²³	Neural Network	71.4	90.0			
Ortiz J. (1995) ²⁴	LDA	67.5	67.4			
Negassa A. (2021) ²⁵	Super Learner			0.83 [0.80-0.87]		

			RF	88.70			
			LightGBM	89.12	0.93		0.071
			KNN	88.28	0.44		0.000
Jahangiri S. (2024) ³⁰			LR	58		0.57	
			RF	60		0.53	
Allam A. (2019) ³¹			Neural Network			0.64 [0.640–0.645]	
			LASSO			0.643 [0.64–0.646]	
Beecy N. A. (2017) ¹⁰			XGBoost (combinado)		0.756	0.756	
			GBM	89	0.90	0.90	
Ju C. (2021) ³²			Decision Tree			0.87	
			LR			0.86	
			LR	38.1		0.580	
Ru B. (2023) ³³			MLP NN	38.5		0.583	
			RF	40.0		0.577	
			XGBoost	51.3		0.595	
Soliman A. (2023) ³⁴			LSTM	78		0.65	0.58
			CatBoost	83		0.66	0.60
			LR	32.4	0.652	0.63 [0.59–0.64]	
			CART	48.6	0.622	0.59 [0.48–0.70]	
Chen S. (2023) ³⁵			XGBoost	38.7	0.589	0.54 [0.38–0.70]	
			Naive Bayes	61.7	0.577	0.58 [0.55–0.61]	
			SVM	18.9	0.659	0.56 [0.18–0.93]	
			RF	29.3	0.649	0.57 [0.29–0.85]	
Ashfaq A. (2019) ³⁶			LSTM			0.77	0.51
			LR		0.66 [0.62–0.69]	0.73 [0.66–0.80]	
			LASSO		0.65 [0.61–0.70]	0.73 [0.67–0.79]	
Angraal S. (2019) ³⁷			RF		0.72 [0.69–0.75]	0.76 [0.71–0.81]	
			GBM		0.68 [0.66–0.71]	0.73 [0.69–0.77]	
			SVM Kernel		0.66 [0.60–0.72]	0.72 [0.63–0.81]	
			SVM Basis		0.65 [0.59–0.72]	0.72 [0.65–0.79]	
			LR				
			Naive Bayes	68.33	72.44		
			RF	82.41	72.63		
Xu C. (2023) ³⁸			Extreme Gradient	99.44	78.96	0.82	
			Boost	93.38	77.33		
			KNN	71.04	62.28		
			Decision Tree	71.52	72.43		
König S (2021) ³⁹			GLM			0.82 [0.81–0.84]	
			RF			0.87 [0.86–0.88]	

	Weighted DT			0.528	0.528
	SVM			0.528	0.528
	Weighted SVM			0.535	0.535
	MLP			0.628	0.628
Wang Z. (2021) ⁴⁶	MVE-EK		79.67	0.896	
	RF			0.76	
Jing L. (2020) ⁴⁷	LR			0.74	
	XGBoost			0.77	
	SVM	62	0.90		
	LS-SVM	86	0.60		
Najafi-Vosough R. ⁴⁸ (2021)	Bagging	46	0.87		
	Ada Boost	84	0.84		
	RF	69	0.90		
	Naive Bayes	66	0.86		
	Boruta-SVM	50	0.81		0.73
Sabouri M. (2023) ⁴⁹	MRMR-LR	63	0.68		0.71
	MRMR-Naive Bayes	54	0.70		0.70
	Boruta-Naive Bayes	50	0.67		0.69

Fonte: próprio autor

Legenda: TAN - Tree Augmented Naive Bayes; GBM - Gradient Boosting Machines; LASSO - Least Absolute Shrinkage and Selection Operator; LR – Logistic Regression; MLP - Multilayer Perceptron; RF - Random Forest; XGBoost - Extreme Gradient Boosting; SVM - Support Vector Machine; MARS - Multivariate Adaptive Regression Splines; Linear SVC - Linear Support Vector Classifier; KNN - K-Nearest Neighbors, LSTM - Long Short-Term Memory; CART - Classification and Regression Trees; GLM - Generalized Linear Model; NNET - Neural Network; LDA - Linear Discriminant Analysis; GNB - Gaussian Naive Bayes; ANN - Artificial Neural Network, LS-SVM - Least Squares Support Vector Machine; FRDLS - Feature Rearrangement based Deep Learning System; MARKER-HF - Machine learning Assessment of Risk and Early mortality in Heart Failure; ACD-VC - Administrative Claim Data-based Voting Classifier; DLS-MSM - Deep Learning System based on Multi-head Self-attention Mechanism.

Os estudos que avaliaram o desfecho de mortalidade, apresentaram uma média para AUC de 0.73, mediana 0.76 (IQR: 0.628-0.852). Para desfecho de readmissão, a média da AUC resultou em 0.64, mediana 0.648 (IQR: 0.595-0.707). Os valores isolados da acurácia, sensibilidade, F-score e AUC estão representados na tabela 4.

5.5 Comparativo com Escores Clássicos

Apenas 9 estudos realizaram o comparativo com escores clássicos. Analisando sob uma perspectiva ampla, 6 desses comparativos tiveram como melhor resultado a predição realizada por IA ou ML, sobretudo três que se destacaram nas médias obtidas da AUC: Greenberg, Tohyama, Beecy e Kwon (tabela 5). Vale ressaltar que no estudo de Greenberg, os autores utilizaram a fração de ejeção do ventrículo esquerdo (LVEF) como parâmetro comparativo, isto é, embora não seja propriamente um escore que avalia a predição de mortalidade ou readmissão, esses autores optaram por seguir com essa análise utilizando os valores da fração de ejeção para prever o desfecho.

Tabela 5 – Comparativo dos modelos de ML e IA com Escores Clássicos

Autor	Escore clássico	AUC M.	AUC R.	IA ou ML	AUC M.	AUC R.
Greenberg B. (2021) ¹⁸	LVEF	0.520		MARKER-HF	0.850	
				Logistic lasso	0.770	
Tohyama T. (2021) ²¹	SHFM MAGGIC	0.713 0.726		SVM	0.760	
				RF	0.760	
				GBT	0.760	
				ACD-VC	0.770	
				Neural Network	0.770	
				LASSO	0.740	
Austin D.E. (2022) ²²	EHMGR	0.755		Neural Network	0.680	
				RF	0.740	
				XGBoost	0.750	
				XGBoost		0.654
Sharma V. (2022) ²⁶	LACE		0.570	GBM		0.650
				Ada Boost		0.646
				CatBoost		0.642
				LightGBM		0.641
				Linear SVC		0.639
				GNB		0.624
				RF		0.617
				Decision Tree		0.597
				LR		0.596
				Neural Network		0.578
Beecy N. A. (2017) ¹⁰	HOSPITAL		0.666	XGBoost	0.756	0.756

Soliman A. (2023) ³⁴	LACE		0.560	LSTM		0.650
				CatBoost		0.660
Angraal S. (2019) ³⁷	KCCQ	0.620		LR	0.660	0.730
				LASSO	0.650	0.730
				RF	0.720	0.760
				GBM	0.680	0.730
				SVM Kernel	0.660	0.720
				SVM Basis	0.650	0.720
				Deep learning	0.782	
Kwon J. (2019) ¹⁵	MAGGIC	0.718		RF	0.696	
				LR	0.703	
				SVM	0.646	
				Bayesian network	0.729	
				LR	0.576	0.576
Awan SE (2019) ⁴⁵	LACE	0.551	0.551	RF	0.501	0.501
				Weighted RF	0.548	0.548
				Decision Trees	0.520	0.520
				Weighted DT	0.528	0.528
				SVM	0.528	0.528
				Weighted SVM	0.535	0.535
				MLP	0.628	0.628

Fonte: próprio autor.

Legenda: AUC M. – Área da Curva ROC Mortalidade; AUC M. – Área da Curva ROC Readmissão; LVEF - Left Ventricular Ejection Fraction; SHFM - Seattle Heart Failure Model; MAGGIC - Meta-Analysis Global Group in Chronic Heart Failure; EHMGR - Emergency Heart Failure Mortality Risk Grade; LACE - Length of stay, Acuity of admission, Charlson comorbidity index, and number of Emergency department visits; KCCQ - Kansas City Cardiomyopathy Questionnaire.

6 DISCUSSÃO

Esta revisão mostrou que sob uma perspectiva ampla, o desempenho médio das machine learning utilizadas não superou as expectativas, equiparando-se ou ficando atrás dos famosos escores clássicos, com ressalvas para alguns modelos que obtiveram um resultado promissor. Nessa perspectiva, com os resultados dessa revisão sistemática, objetiva-se analisar o desempenho alcançado pelas machine learning ou inteligência artificial ao redor do mundo quanto a eficácia na predição de prognóstico em pacientes com insuficiência cardíaca, assim como avaliar sua capacidade em superar os principais escores utilizados na atualidade. Isso porque a capacidade dos escores clássicos muitas vezes é insuficiente, uma vez que não estão em constante atualizações, e quando há, depende da busca constante dos profissionais de saúde pelos famosos *guidelines*. Dessa forma, esta revisão fornece detalhes importante sobre como devemos abordar a inteligência artificial nesse cenário.

Hodiernamente, a maioria dos escores são criados com intuito de avaliar a mortalidade, não tendo muito parâmetros que ajudem na predição de readmissão. Por conta disso, vale salientar ainda, que ao comparar os desfechos analisados, o desempenho das máquinas em prever a readmissão dos pacientes foi inferior comparado a mortalidade, com diferença aproximada de 0.1 na AUC média. Dito isso, a predição de mortalidade se mostrou mais consistente, já que as variáveis envolvidas nesse tipo de previsão são mais objetivas e causam menos confusão nas análises.

Ainda assim, mesmo com as limitações dos escores clássico, sobretudo quando se trata das populações restritas e que não reflete a população em geral – um dos maiores empecilhos desse tipo de escore – observa-se que de maneira geral os modelos de ML não superaram significativamente esses escores, como por exemplo Seattle Heart Failure Model, com AUC para mortalidade de 0,729 (IC 95%, 0,714 a 0,744); Metabolic Exercise Cardiac Kidney Index (MECKI) com AUC 0,804 (0,754-0,852); Heart Survival Score com AUC de 0,76; Emergency Heart Failure Mortality Risk Grade (EHMGR) com AUC de 0.755; MAGGIC Score com AUC de 0,718.^{6,50,51}

Contudo, saindo um pouco da perspectiva ampla e da inferência obtida através da média dos resultados, podemos perceber que algumas machines learning superaram as expectativas e se destacaram dos escores clássicos, provando seu potencial em

situação de predição. Apenas 25% dos artigos realizaram um comparativo com esses escores, todavia, é possível notar que desses contrapostos, considerando a média do desempenho das IA e ML analisadas em cada estudo, a maioria dos modelos alcançaram resultados superiores aos escores clássicos.

Então, surge o questionamento do porquê essas ML em especiais conseguiram alcançar esses resultados enquanto outras não conseguiram nem se equiparar aos escores. Isso acontece porque além dos modelos convencionais de inteligência artificial, há subprodutos com abordagens e processamento de dados diferentes que alguns dos autores utilizaram na tentativa de mapear melhor as variáveis a fim de que a máquina consiga construir melhor as redes neurais e árvores de dados, pois são estes diferenciais fornecidos pela IA e ML.

Outrossim, o excesso de variáveis advinda das fontes de dados podem ser tanto um fator de melhora, como um fator de piora para o desempenho da máquina em predizer os desfechos, em razão de uma má escolha das variáveis que pode resultar em redes neurais e as ramificações anômalas formadas pelo processamento dos dados e que confundem a inteligência artificial. Em cenários como esses, é preciso selecionar adequadamente as variantes e começar do zero o processamento, posto que as redes neurais precisam ser formadas desde o início.

Outro limitador importante que podemos observar, é a falta de informações clínicas significativas para a população alvo, como por exemplo a classificação funcional dos pacientes (New York Heart Association – NYHA) que podem estar faltando nos prontuários eletrônicos, sobretudo nos registros médicos, em que muitas vezes os pacientes são atendidos, mas o profissional não anota esse tipo de dado, comprometendo mais ainda as ramificações das redes neurais. Somado a isso, temos a limitação geográfica dos estudos, em que a maioria foi realizada em países considerados de primeiro mundo, com população majoritariamente branca e características demográficas que podem não refletir a variação etiológica da insuficiência cardíaca. Portanto, é preciso ter cuidado ao afirmar que essas máquinas são excelentes para predição. Talvez se alguns pesquisadores focassem em alimentar as máquinas com dados mais significativos e sem dados faltantes relativos a IC, os resultados se manteriam consistente e com maior acurácia. Não obstante disso, por se tratar de uma revisão sistemática realizada apenas por um único revisor,

devemos ter cuidado quanto a representatividade dos artigos incluídos, já que pode existir um viés de seleção que compromete a análise e resultados dos artigos.

Embora haja essas limitações, nos últimos 5 anos, é inegável a crescente utilização de inteligência artificial ou machine learning como ferramentas facilitadoras. Segundo a lei de Moore, a indústria tecnológica dobra a cada ano sua capacidade de processamento de dados realizado pelas placas de vídeo e processadores. Nesse contexto, a indústria tenta criar modelos importantes para sociedade, na expectativa de que essas superem a capacidade humana em realizar tarefas. No âmbito da medicina não é diferente, diversas startups correm para dominar o cenário da saúde, propondo fornecer máquinas capazes de auxiliarem o trabalho do profissional de saúde, bem como avaliar o paciente tanto da perspectiva diagnóstica, como prognóstica.

Em virtude dessa conjuntura, através dessa revisão podemos perceber que há um potencial bastante escalável para esses modelos de predição de prognóstico. Se tratando de uma evolução tecnológica relativamente recente, essa revisão mostrou também a importância e possibilidade de utilizarmos ferramentas tecnológicas para além do lazer, mas também como promoção e prevenção no âmbito da saúde. Nessa perspectiva, ainda conseguimos contribuir com a discussão sobre a eficácia dessas máquinas, mitigando a desconfiança pelos profissionais da saúde, e incentivando-os a usufruir dessas técnicas nas práticas clínicas e hospitalar.

7 CONCLUSÃO

Indubitavelmente, conclui-se que a utilização de machine learning e inteligência artificial nesse cenário não tem mais volta. Embora muitos desses modelos ainda se equiparam aos escores clássicos, há outros que se destacam. É preciso ter cautela ao treinar esses algoritmos, uma vez que há muitas variáveis que precisam ser controladas, dificultando o processo de aprendizagem da máquina e consequentemente comprometendo uma boa performance, sobretudo na predição de prognóstico.

Não obstante, esta revisão reforçou que a comunidade científica, sobretudo voltada para âmbito da saúde, está no caminho certo. Notáveis resultados foram alcançados

até agora e desmitificaram a ideia de que IA e ML não pode ser usada na prática clínica e hospitalar. Corroborou para difundir essas máquinas como ferramentas auxiliaadoras e de grande aplicabilidade, possibilitando uma ação multidisciplinar mais enfática nos grupos com piores prognósticos previstos, não substituindo os profissionais de saúde, mas otimizando seu trabalho em prol do bem-estar do paciente.

REFERÊNCIAS:

1. Rohde LEP, Montera MW, Bocchi EA, Clausell NO, de Albuquerque DC, Rassi S, et al. Diretriz brasileira de insuficiência cardíaca crônica e aguda. *Arq Bras Cardiol*. 2018 Sep 1;111(3):436–539.
2. Nogueira PR, Rassi S, De Corrêa KS. Epidemiological, clinical e therapeutic profile of heart failure in a tertiary hospital. *Arq Bras Cardiol* [Internet]. 2010 Sep [cited 2023 Oct 16];95(3):392–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/20949226/>
3. Instituto Brasileiro de Geografia e Estatística. População cresce, mas número de pessoas com menos de 30 anos cai 5,4% de 2012 a 2021 | Agência de Notícias [Internet]. Agência IBGE Notícias. 2022 [cited 2023 Oct 16]. Available from: <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/34438-populacao-cresce-mas-numero-de-pessoas-com-menos-de-30-anos-cai-5-4-de-2012-a-2021>
4. Eustáquio J, Alves D, De Janeiro R. DEMOGRAFIA E ECONOMIA Nos 200 anos da Independência do Brasil e cenários para o século XXI. 2022;
5. Beam AL, Drazen JM, Kohane IS, Leong TY, Manrai AK, Rubin EJ. Artificial Intelligence in Medicine. <https://doi.org/101056/NEJMe2206291> [Internet]. 2023 Mar 29 [cited 2023 Oct 29];388(13):1220–1. Available from: <https://www.nejm.org/doi/full/10.1056/NEJMe2206291>
6. Levy WC, Mozaffarian D, Linker DT, Sutradhar SC, Anker SD, Cropp AB, et al. The Seattle Heart Failure Model: Prediction of survival in heart failure. *Circulation* [Internet]. 2006 Mar [cited 2023 Nov 21];113(11):1424–33. Available from: <http://www.circulationaha.org>
7. Rich JD, Burns J, Freed BH, Maurer MS, Burkhoff D, Shah SJ. Meta-analysis global group in chronic (MAGGIC) heart failure risk score: Validation of a simple tool for the prediction of morbidity and mortality in heart failure with preserved ejection fraction. *J Am Heart Assoc* [Internet]. 2018 Oct 1 [cited 2023 Nov 21];7(20). Available from: <http://www.clinicaltrials.gov>.
8. Agostoni P, Corrà U, Cattadori G, Veglia F, La Gioia R, Scardovi AB, et al. Metabolic exercise test data combined with cardiac and kidney indexes, the MECKI score: a multiparametric approach to heart failure prognosis. *Int J Cardiol* [Internet]. 2013 [cited 2023 Nov 21];167(6):2710–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/22795401/>
9. Haug CJ, Drazen JM. Artificial Intelligence and Machine Learning in Clinical Medicine, 2023. *New England Journal of Medicine* [Internet]. 2023 Mar 30 [cited 2023 Nov 21];388(13):1201–8. Available from: <https://www.nejm.org/doi/full/10.1056/NEJMra2302038>
10. Beecy AN, Gummalla M, Sholle E, Xu Z, Zhang Y, Michalak K, et al. Utilizing electronic health data and machine learning for the prediction of 30-day unplanned readmission or all-cause mortality in heart failure. *Cardiovasc Digit Health J*. 2020 Sep 1;1(2):71–9.
11. Chicco D, Jurman G. Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC Med Inform Decis Mak* [Internet]. 2020 Feb 3 [cited 2024 Apr 21];20(1):1–16. Available from: <https://bmcmmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-020-1023-5>

12. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. *New England Journal of Medicine* [Internet]. 2019 Apr 4 [cited 2024 Apr 21];380(14):1347–58. Available from: <https://www.nejm.org/doi/full/10.1056/NEJMra1814259>
13. Sci-Hub | Machine learning in cardiac CT: Basic concepts and contemporary data. *Journal of Cardiovascular Computed Tomography*, 12(3), 192–201 | 10.1016/j.jcct.2018.04.010 [Internet]. [cited 2024 Apr 28]. Available from: <https://sci-hub.se/https://doi.org/10.1016/j.jcct.2018.04.010>
14. Krittanawong C, Virk HUH, Bangalore S, Wang Z, Johnson KW, Pinotti R, et al. Machine learning prediction in cardiovascular diseases: a meta-analysis. *Sci Rep* [Internet]. 2020 Dec 1 [cited 2024 Apr 28];10(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/32994452/>
15. Kwon JM, Kim KH, Jeon KH, Lee SE, Lee HY, Cho HJ, et al. Artificial intelligence algorithm for predicting mortality of patients with acute heart failure. *PLoS One*. 2019 Jul 1;14(7).
16. Frizzell JD, Liang L, Schulte PJ, Yancy CW, Heidenreich PA, Hernandez AF, et al. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: Comparison of machine learning and other statistical approaches. *JAMA Cardiol*. 2017 Feb 1;2(2):204–9.
17. Wang Z, Chen L, Zhang J, Yin Y, Li D. Multi-view ensemble learning with empirical kernel for heart failure mortality prediction. *Int J Numer Method Biomed Eng*. 2020 Jan 1;36(1).
18. Greenberg B, Adler E, Campagnari C, Yagil A. A machine learning risk score predicts mortality across the spectrum of left ventricular ejection fraction. *Eur J Heart Fail*. 2021 Jun 1;23(6):995–9.
19. Mahajan SM, Mahajan AS, King R, Negahban S. Predicting risk of 30-day readmissions using two emerging machine learning methods. In: *Studies in Health Technology and Informatics*. IOS Press; 2018. p. 250–5.
20. Tian P, Liang L, Zhao X, Huang B, Feng J, Huang L, et al. Machine Learning for Mortality Prediction in Patients With Heart Failure With Mildly Reduced Ejection Fraction. *J Am Heart Assoc*. 2023 Jun 20;12(12).
21. Tohyama T, Ide T, Ikeda M, Kaku H, Enzan N, Matsushima S, et al. Machine learning-based model for predicting 1 year mortality of hospitalized patients with heart failure. *ESC Heart Fail*. 2021 Oct 1;8(5):4077–85.
22. Austin DE, Lee DS, Wang CX, Ma S, Wang X, Porter J, et al. Comparison of machine learning and the regression-based EHMRG model for predicting early mortality in acute heart failure. *Int J Cardiol*. 2022 Oct 15;365:78–84.
23. Adler ED, Voors AA, Klein L, Macheret F, Braun OO, Urey MA, et al. Improving risk prediction in heart failure using machine learning. *Eur J Heart Fail*. 2020 Jan 1;22(1):139–47.
24. Ortiz J, M Ghefter CG, S Silva CE, E Sabbatini RM. One-Year Mortality Prognosis in Heart Failure: A Neural Network Approach Based on Echocardiographic Data. Vol. 26, *JACC*. 1995.
25. Negassa A, Ahmed S, Zolty R, Patel SR. Prediction Model Using Machine Learning for Mortality in Patients with Heart Failure. *American Journal of Cardiology*. 2021 Aug 15;153:86–93.

26. Sharma V, KULKARNI V, MCALISTER F, EURICH DEAN, KESHWANI S, SIMPSON SH, et al. Predicting 30-Day Readmissions in Patients With Heart Failure Using Administrative Data: A Machine Learning Approach. *J Card Fail.* 2022 May 1;28(5):710–22.
27. Mahajan SM, Ghani R. Using ensemble machine learning methods for predicting risk of readmission for heart failure. In: *Studies in Health Technology and Informatics.* IOS Press; 2019. p. 243–7.
28. Gutman R, Aronson D, Caspi O, Shalit U. What drives performance in machine learning models for predicting heart failure outcome? *European Heart Journal - Digital Health.* 2023 May 1;4(3):175–87.
29. Li D, Fu J, Zhao J, Qin J, Zhang L. A deep learning system for heart failure mortality prediction. *PLoS One.* 2023 Feb 1;18(2 February).
30. Jahangiri S, Abdollahi M, Rashedi E, Azadeh-Fard N. A machine learning model to predict heart failure readmission: toward optimal feature set. *Front Artif Intell.* 2024;7.
31. Allam A, Nagy M, Thoma G, Krauthammer M. Neural networks versus Logistic regression for 30 days all-cause readmission prediction. *Sci Rep.* 2019 Dec 1;9(1).
32. Ju C, Zhou J, Lee S, Tan MS, Liu T, Bazoukis G, et al. Derivation of an electronic frailty index for predicting short-term mortality in heart failure: a machine learning approach. *ESC Heart Fail.* 2021 Aug 1;8(4):2837–45.
33. Ru B, Tan X, Liu Y, Kannapur K, Ramanan D, Kessler G, et al. Comparison of Machine Learning Algorithms for Predicting Hospital Readmissions and Worsening Heart Failure Events in Patients With Heart Failure With Reduced Ejection Fraction: Modeling Study. *JMIR Form Res.* 2023;7.
34. Soliman A, Agvall B, Etminani K, Hamed O, Lingman M. The Price of Explainability in Machine Learning Models for 100-Day Readmission Prediction in Heart Failure: Retrospective, Comparative, Machine Learning Study. *J Med Internet Res.* 2023;25.
35. Chen S, Hu W, Yang Y, Cai J, Luo Y, Gong L, et al. Predicting Six-Month Re-Admission Risk in Heart Failure Patients Using Multiple Machine Learning Methods: A Study Based on the Chinese Heart Failure Population Database. *J Clin Med.* 2023 Feb 1;12(3).
36. Ashfaq A, Sant'Anna A, Lingman M, Nowaczyk S. Readmission prediction using deep learning on electronic health records. *J Biomed Inform.* 2019 Sep 1;97.
37. Angraal S, Mortazavi BJ, Gupta A, Khera R, Ahmad T, Desai NR, et al. Machine Learning Prediction of Mortality and Hospitalization in Heart Failure With Preserved Ejection Fraction. *JACC Heart Fail.* 2020 Jan 1;8(1):12–21.
38. Xu C, Li H, Yang J, Peng Y, Cai H, Zhou J, et al. Interpretable prediction of 3-year all-cause mortality in patients with chronic heart failure based on machine learning. *BMC Med Inform Decis Mak.* 2023 Dec 1;23(1).
39. König S, Pellissier V, Hohenstein S, Bernal A, Ueberham L, Meier-Hellmann A, et al. Machine learning algorithms for claims data-based prediction of in-hospital mortality in patients with heart failure. *ESC Heart Fail.* 2021 Aug 1;8(4):3026–36.

40. Ketabi M, Andishgar A, Fereidouni Z, Sani MM, Abdollahi A, Vali M, et al. Predicting the risk of mortality and rehospitalization in heart failure patients: A retrospective cohort study by machine learning approach. *Clin Cardiol.* 2024 Feb 1;47(2).
41. Shameer K, Johnson KW, Yahi A, Miotto R, Li LI, Ricks D, et al. Icahn Institute of Genomics and Multiscale Biology 5. Zena and Michael A. Wiener Cardiovascular Institute, Icahn School of Medicine at Mount Sinai 6. Department of Anesthesiology, Icahn School of Medicine at Population Health Science and Policy [Internet]. Available from: www.worldscientific.com
42. Lv H, Yang X, Wang B, Wang S, Du X, Tan Q, et al. Machine learning-driven models to predict prognostic outcomes in patients hospitalized with heart failure using electronic health records: Retrospective study. *J Med Internet Res.* 2021 Apr 1;23(4).
43. Sarijaloo FB, Park J, Zhong X, Wokhlu A. Predicting 90 day acute heart failure readmission and death using machine learning-supported decision analysis. *Clin Cardiol.* 2021 Feb 1;44(2):230–7.
44. van der Galiën OP, Hoekstra RC, Gürgöze MT, Manintveld OC, van den Bunt MR, Veenman CJ, et al. Prediction of long-term hospitalisation and all-cause mortality in patients with chronic heart failure on Dutch claims data: a machine learning approach. *BMC Med Inform Decis Mak.* 2021 Dec 1;21(1).
45. Awan SE, Bennamoun M, Sohel F, Sanfilippo FM, Dwivedi G. Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics. *ESC Heart Fail.* 2019 Apr 1;6(2):428–35.
46. Wang Z, Zhu Y, Li D, Yin Y, Zhang J. Feature rearrangement based deep learning system for predicting heart failure mortality. *Comput Methods Programs Biomed.* 2020 Jul 1;191.
47. Jing L, Ulloa Cerna AE, Good CW, Sauers NM, Schneider G, Hartzel DN, et al. A Machine Learning Approach to Management of Heart Failure Populations. *JACC Heart Fail.* 2020 Jul 1;8(7):578–87.
48. Najafi-Vosough R, Faradmaj J, Hosseini SK, Moghimbeigi A, Mahjub H. Predicting hospital readmission in heart failure patients in Iran: A comparison of various machine learning methods. *Healthc Inform Res.* 2021 Oct 1;27(4):307–14.
49. Sabouri M, Rajabi AB, Hajianfar G, Gharibi O, Mohebi M, Avval AH, et al. Machine learning based readmission and mortality prediction in heart failure patients. *Sci Rep.* 2023 Dec 1;13(1).
50. Adamopoulos S, Miliopoulos D, Piotrowicz E, Snoek JA, Panagopoulou N, Nanas S, et al. International validation of the Metabolic Exercise test data combined with Cardiac and Kidney Indexes (MECKI) score in heart failure. *Eur J Prev Cardiol* [Internet]. 2023 Sep 20 [cited 2024 Apr 28];30(13):1371–9. Available from: <https://dx.doi.org/10.1093/eurjpc/zwad191>
51. Qian S, Cao B, Li P, Dong N. Development and validation of mortality prediction models for heart transplantation using nutrition-related indicators: a single-center study from China. *Front Cardiovasc Med.* 2024 Feb 26;11:1346202.