

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ
ESCOLA DE MEDICINA E CIÊNCIAS DA VIDA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA SAÚDE

**PREDIÇÃO DE MORTALIDADE E REINTERNAMENTO EM 48 HORAS EM
UNIDADES DE TERAPIA INTENSIVA UTILIZANDO MODELOS DE *MACHINE*
*LEARNING***

Aluno: Gustavo Henrique dos Santos Silva
Orientadora: Prof.^a Dr.^a Cristina Pelegrino Baena
Co-orientador: Prof.^a Dr. Murilo Guedes

Curitiba, novembro de 2023

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ
ESCOLA DE MEDICINA E CIÊNCIAS DA VIDA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA SAÚDE

**PREDIÇÃO DE MORTALIDADE E REINTERNAMENTO EM 48 HORAS EM
UNIDADES DE TERAPIA INTENSIVA UTILIZANDO MODELOS DE *MACHINE*
*LEARNING***

Dissertação de mestrado acadêmico apresentada ao Programa de Pós-graduação em Ciências da Saúde como requisito parcial para obtenção do título de Mestre em Ciências da Saúde.

Orientadora: Prof.^a Dr.^a Cristina Pelegrino Baena

Co-orientador: Prof.^a Dr. Murilo Henrique Guedes

Curitiba, novembro de 2023



Pontifícia Universidade Católica do Paraná
Escola de Medicina e Ciências da Vida
Programa de Pós-Graduação em Ciências da Saúde

ATA DA SESSÃO PÚBLICA DE EXAME DE DISSERTAÇÃO DO PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA SAÚDE EM NÍVEL DE MESTRADO DA PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ.

Aos 08 dias do mês de dezembro de 2023 às 13:00, realizou-se a sessão pública de Defesa de Dissertação "PREDIÇÃO DE MORTALIDADE E REINTERNO EM 48 HORAS EM UNIDADES DE TERAPIA INTENSIVA UTILIZANDO MODELOS DE MACHINE LEARNING" apresentado por Gustavo Henrique dos Santos Silva para obtenção do título de mestre; Área de concentração: Pesquisa Médica Translacional.

A banca examinadora foi composta pelos seguintes membros:

MEMBROS DA BANCA	ASSINATURA
Profa. Dra. Cristina Pellegrino Baena - Presidente	
Prof. Dr. Felipe Francisco Bondan Tuon - PUCPR	
Prof. Dr. Juliano Gasparetto - PUCPR	
Profa. Dra. Cristina Terumi Okamoto - UP	

De acordo com as normas regimentais a Banca Examinadora deliberou sobre os conceitos a serem distribuídos e que foram os seguintes:

Profa. Dra. Cristina Pellegrino Baena
Prof. Dr. Felipe Francisco Bondan Tuon
Prof. Dr. Juliano Gasparetto
Profa. Dra. Cristina Terumi Okamoto

Conceito: *Aprovado*
Conceito: *APROVADO*
Conceito: *Aprovado*
Conceito: *Aprovado*

Observações da Banca Examinadora:

Profa. Dra. Cristina Pellegrino Baena
Presidente da Banca Examinadora

Profa. Dra. Cristina Pellegrino Baena
Coordenadora do PPGCS-PUCPR

AGRADECIMENTOS

Gostaria, primeiramente, de agradecer a Pontifícia Universidade Católica do Paraná pela oportunidade de lecionar durante os últimos 2 anos e pela oportunidade de iniciar meus estudos em pesquisa científica por meio desta dissertação de mestrado. Em especial ao Dr. Juliano Gasparetto por me abrir as portas para o Hospital Marcelino Champagnat e a Universidade. Também gostaria de deixar registrado o meu agradecimento a minha orientadora, Dra. Cristina Pellegrino Baena, por todo o esforço em fazer este projeto acontecer, por toda a paciência e apoio durante o processo de construção desta dissertação.

Deixo também um agradecimento a equipe da MunaiTM, principalmente ao Dr. Murilo por todo o suporte e orientações e Augusto, Bruno e Hugo, por todo o suporte fornecido com dados e estatísticas para que esse trabalho acontecesse, por todas as noites e finais de semana que trabalharam para me fornecer mais dados e responder meus questionamentos.

Gostaria de agradecer, especialmente, a minha esposa, Raíssa Campos D'Amico, por todo o apoio durante esses últimos 4 anos, por me incentivar diariamente a ser o meu melhor e por me lembrar em muitas ocasiões que eu sou capaz de realizar muito mais do que eu imagino que posso. Obrigado por sonhar acordada comigo e ajudar a transformar nossos sonhos em realidade.

Mais importante, gostaria de agradecer aos meus pais, por todo o suporte e educação ao longo de toda a minha vida. Sei que não foi nada fácil para dois fisioterapeutas bancarem escola particular, cursinho e faculdade particular de medicina para dois filhos. Enquanto escrevo este texto lembro de quando minha mãe ou meu pai passavam me buscar na escola e me levavam para o hospital ou para a clínica, onde eu ficava sentado fazendo minha tarefa

de casa enquanto eles faziam o terceiro turno do dia. Educação sempre foi uma prioridade na vida dos meus pais e só posso ser eternamente grato por todo o esforço que fizeram para que eu chegasse até aqui. E não poderia terminar esses agradecimentos sem lembrar da minha avó Edna, falecida em agosto desse ano, professora aposentada, que fazia questão de deixar na sala de sua casa a foto da formatura de todos os filhos e netos que terminaram a faculdade. Tenho uma base muito forte e isso vem de casa. Obrigado, família amada.

LISTA DE SIGLAS E ABREVIATURAS

APACHE II - *acute physiology and chronic health evaluation*

AUC *Area under curve* – área sob a curva

CONEP Comitê Nacional de Ética em Pesquisa

DP Desvio padrão

EAP Escores de Alerta Precoce

FIGS *Fast Interpretable Freedy-Tree Sums*

IA Inteligência Artificial

LGBM *Light Gradient Boosting Machine*

MEWS *Modified Early Warning Score*

ML *Machine Learning*

NEWS *National Early Warning Score*

qSOFA *Quick Sequential Organ Failure Assessment*

SSDC Sistema de suporte de decisão clínica

SOFA *Sequential organ failure assessment*

ROC *Receiver Operator Characteristic*

UTI Unidade de Terapia intensiva

LISTA DE FIGURAS

Figura 1. Janela de tempo de dados avaliados.....	8
Figura 2. Fluxograma dos dados extraídos.....	12
Figura 3. Curva ROC para os diferentes modelos de <i>machine learning</i>	16
Figura 4. Peso atribuído para as variáveis utilizando o modelo LGBM	17

LISTA DE TABELAS

Tabela 1. Variáveis utilizadas para treinamento dos modelos de machine learning	8
Tabela 2. Características demográficas	13
Tabela 3. Métricas avaliadas nos testes para os diferentes modelos de ML e MEWS.....	14
Tabela 4. Diferenças entre os pacientes excluídos do estudo	15

RESUMO

INTRODUÇÃO: A readmissão na Unidade de Terapia Intensiva (UTI) levanta preocupações significativas, uma vez que pacientes readmitidos em UTI apresentam taxas de mortalidade mais elevadas, maior tempo de internamento hospitalar e maiores custos com assistência médica, quando comparados com pacientes que não são readmitidos em UTI após a sua alta. Para enfrentar esse desafio, modelos de *machine learning* surgiram como ferramentas potenciais para auxiliar os médicos na tomada de decisões de alta na beira do leito. Modelos que utilizam dados do paciente, como sinais vitais e outras variáveis, podem ser treinados para prever o risco de readmissão ou morte nas primeiras 48 horas.

MÉTODOS: Neste estudo, foram utilizados dados clínicos de pacientes internados em 13 hospitais brasileiros. Modelos de *machine learning* foram treinados utilizando os últimos cinco sinais vitais e variáveis do paciente e do internamento (como idade, sexo, tempo de internamento, plano de saúde e especialidade médica) para análise do risco de readmissão ou morte dentro de 48 horas após a alta da UTI. Para a análise do desfecho composto em questão foram utilizados os modelos de *machine learning* *Light Gradient Boosting Model* (LGBM), *TabNet*, *RuleFit*, *Fast Interpretable Greedy-Tree Sums* (FIGS) e o escore *Modified Early Warning Score* (MEWS) e foram avaliadas, para todos os modelos, as métricas de área sob a curva ROC (AUC), acurácia, precisão, escore F1, precisão média e recall.

RESULTADOS: A população foi composta por 11.948 pacientes únicos, entre os quais houve 1.924 óbitos e 2.030 readmissões na UTI nas primeiras 48 horas após a alta. O modelo LGBM exibiu uma acurácia de 91,3% e uma AUC de 0,82, o *TabNet* registrou uma acurácia de 84,7% e uma AUC de 0,81, o *RuleFit* apresentou uma acurácia de 77,9% e uma AUC de 0,79, o FIGS obteve uma acurácia de 90,2% e uma AUC de 0,78, enquanto o MEWS demonstrou uma precisão de 89,1% e uma AUC de 0,75.

CONCLUSÃO: Os resultados do estudo revelam a superioridade dos modelos de *machine learning* em relação às pontuações tradicionais de alerta precoce. Esses modelos podem servir como valiosas ferramentas na beira do leito para os médicos ao decidir sobre alta da UTI. Os modelos de *machine learning* apresentam resultados promissores na previsão do risco de readmissão ou morte nas primeiras 48 horas após a alta da UTI.

PALAVRAS-CHAVE: *Machine learning*, unidade de terapia intensiva, inteligência artificial, mortalidade.

ABSTRACT

INTRODUCTION: Readmission to the Intensive Care Unit (ICU) raises significant concerns, as patients readmitted to the ICU have higher mortality rates, longer hospital stays, and increased healthcare costs compared to patients who are not readmitted to the ICU after their discharge. To face this challenge, *machine learning* models have emerged as potential tools to aid physicians in making discharge decisions at the point of care. Models using patient data, such as vital signs and other variables, can be trained to predict the risk of readmission or death within the first 48 hours. **METHODS:** In this study, data from a database comprising 13 Brazilian hospitals was utilized. Machine learning models were trained using the patient's last five vital signs and hospitalization variables (such as age, gender, length of stay, healthcare plan, and medical specialty) to analyze the risk of readmission or death within 48 hours after ICU discharge. For the analysis of the composite outcome in question, the machine learning models Light Gradient Boosting Model (LGBM), TabNet, RuleFit, Fast Interpretable Greedy-Tree Sums (FIGS), and the Modified Early Warning Score (MEWS) were employed. Metrics evaluated for all models included Area Under the ROC Curve (AUC), accuracy, precision, F1 score, average precision, and recall. **RESULTS:** The dataset consisted of 11,948 unique patients, among which there were 1,924 deaths and 2,030 readmissions in the ICU within the first 48 hours after discharge. The LGBM model exhibited an accuracy of 91.3% and an Area Under the Curve (AUC) of 0.82, TabNet recorded an accuracy of 84.7% and an AUC of 0.81, RuleFit showed an accuracy of 77.9% and an AUC of 0.79, FIGS achieved an accuracy of 90.2% and an AUC of 0.78, while MEWS demonstrated an accuracy of 89.1% and an AUC of 0.75. **CONCLUSION:** The study findings reveal the superiority of *machine learning* models over traditional early warning scores. These models can serve as valuable point-of-care tools for physicians when deciding on ICU discharges. *Machine learning* models exhibit promising results in predicting the risk of readmission or death within the initial 48 hours post-ICU discharge.

Keywords: *machine learning*, intensive care units, artificial intelligence, patient discharge, mortality

SUMÁRIO

1. INTRODUÇÃO	1
1.1 EPIDEMIOLOGIA	1
1.2 ESCORES DE ALERTA PRECOCE	2
1.3 INTELIGENCIA ARTIFICIAL	3
1.4 MODELOS DE MACHINE LEARNING	4
1.5 JUSTIFICATIVA	5
2. OBJETIVOS	5
2.1 OBJETIVO GERAL	5
2.2 OBJETIVOS ESPECÍFICOS	5
3. MATERIAIS E MÉTODOS	6
3.1 CRITÉRIOS DE INCLUSÃO E EXCLUSÃO	6
3.2 VARIÁVEIS DE INTERESSE	7
3.3 MODELOS DE MACHINE LEARNING	10
3.4 COMPARAÇÕES DE ACURÁCIA	10
4. RESULTADOS	11
4.1. PARTICIPANTES	11
4.2 DESCRIÇÃO DOS DADOS	12
5. DISCUSSÃO	17
6. CONCLUSÃO	22
7. REFERÊNCIAS	23

1. INTRODUÇÃO

1.1 EPIDEMIOLOGIA

As taxas de readmissão em unidades de terapia intensiva (UTI) tornaram-se um indicador de desempenho dos hospitais e dos profissionais de saúde, bem como um meio de incentivar a prestação eficiente, de alta qualidade e coordenada de cuidados de saúde aos pacientes. (1) O Comitê de Indicadores de Qualidade da Sociedade de Medicina Intensiva recomendou que readmissões dentro de 48 horas sejam um indicador primordial da qualidade dos cuidados prestados na UTI. (2)

A habilidade dos médicos que trabalham em UTI, independentemente de sua experiência ou função, mostram uma precisão moderada na previsão de readmissões à UTI. Isso implica que as capacidades prognósticas baseadas apenas no quadro clínico apresentam limitações. (3)

O cuidado na UTI é responsável por uma grande proporção dos custos hospitalares, sendo estimado que até 40% dos custos hospitalares sejam decorrentes de pacientes críticos e muitas vezes, pacientes são acometidos por eventos adversos que poderiam ser prevenidos, como readmissões após a alta. (4) Até 64% desses eventos acontecem nas primeiras 72 horas após a alta do ambiente de terapia intensiva. As readmissões em UTI são frequentes, com taxas que variam na literatura entre 0,89% (5) e 19% (7), entretanto, nem todos os estudos avaliados demonstram informação acerca das características dos hospitais nos quais foram conduzidos, levantando questionamentos sobre a validade externa dos seus achados. (5) Estes pacientes apresentam maior tempo de internação hospitalar e mortalidade até 10 vezes maior, quando comparados com aqueles pacientes que não foram readmitidos após a alta da UTI, resultando em maiores custos. (7,8). Se a deterioração clínica que acomete esses

pacientes é diagnosticada e manejada de maneira precoce, cerca de 10% das readmissões podem ser prevenidas. (9)

1.2 ESCORES DE ALERTA PRECOCE

Escore de alerta precoce são uma proposta de solução em potencial. Essas ferramentas são modelos de predição clínica que se utilizam de dados vitais para monitorar a saúde dos pacientes durante a hospitalização. Esses modelos identificam a probabilidade de deterioração clínica, geralmente definida como morte ou necessidade de admissão em UTI e podem ser calculados pelos médicos na beira do leito. (10,11)

Os Escores de alerta precoce comumente utilizados são o *Modified Early Warning Score* (MEWS), *National Early Warning Score* (NEWS) e o *Quick Sequential Organ Failure Assessment* (qSOFA). (12,13,14) Esses escores podem ser benéficos para detectar pacientes em alto risco de readmissão na UTI. Entretanto, a sua acurácia pode ser limitada devido à maior complexidade dos casos internados em ambiente de terapia intensiva, quando comparados a enfermarias (8,15). As ferramentas para medir a gravidade do paciente são o *Acute Physiology and Chronic Health Evaluation* (APACHE II) (16) e o *Sequential Organ Failure Assessment* (SOFA) (17). Outra alternativa são programas de transição de cuidados, que tem por objetivo realizar uma redução de cuidados mais lenta para pacientes com maior risco, porém, estudos demonstram que a eficácia dessa estratégia depende de fatores organizacionais. (8,18)

Um sistema de suporte de decisão clínica poderia ser útil para reduzir readmissões já que outros estudos sugerem que mais de 10% das readmissões podem ser prevenidas. (19,20,21) Ferramentas para decisão clínica utilizando dados do paciente em tempo real, na beira do leito, podem fornecer aos médicos informações adicionais para auxiliar na tomada de decisão de indicar alta da UTI para o paciente. (9)

Uma meta-análise (11) que avaliou os escores de alerta precoce concluiu que as respostas clínicas a valores de escores elevados podem apresentar efeitos deletérios e impactar a carga de trabalho. A mudança para a implementação eletrônica dos escores de alerta precoce, entretanto, apresenta uma oportunidade para a introdução de melhores sistemas de escore, particularmente com o aumento do interesse em modelos modernos, como inteligência artificial e *machine learning*. (11)

1.3 INTELIGENCIA ARTIFICIAL

O termo inteligência artificial (IA) data de 1956 e é um braço da engenharia que implementa novos conceitos e novas soluções para problemas complexos. A medicina foi precocemente identificada como uma das áreas mais promissoras para o uso de IA. O seu uso implica em modelar comportamento inteligente em um computador, com a mínima intervenção humana. Sua aplicação aprimorou diagnósticos e decisões em diversas áreas da medicina. Portanto, médicos precisarão se adaptar a essa nova ferramenta. (22,23)

O *machine learning* (ML), é um subtipo da inteligência artificial, onde ocorre uma abordagem baseada em dados. É um programa de computador, que, apesar de não ser diretamente programado, possui a capacidade de aumentar a sua experiência a partir de dados e implementar uma dada tarefa progressivamente melhor, de acordo com indicadores de performance. Os modelos recentemente desenvolvidos superam os tradicionais escores de alerta precoce para a predição de mortalidade. (22-28)

Além disso, ML pode auxiliar decisões clínicas em diversos contextos, como cardiologia, infectologia e em terapia intensiva. (29-32) *Machine learning* também foi avaliada em outros estudos para predição de desfechos clínicos. (9,33,34,35) e pode ser uma ferramenta útil para dar suporte a equipe médica quando ocorre deterioração clínica, reduzindo o tempo para intervenção e tratamento. (36,37)

Machine learning pode proporcionar aos pacientes diagnósticos mais acurados e tratamentos individualizados (38) e demonstrou performance superior para prever desfechos em cuidados críticos e hospitalização quando comparado com o índice de gravidade em emergência. (39)

O impacto desses sistemas nos desfechos dos pacientes, como redução de mortalidade e readmissões em UTI se mantém pouco estudado, devido ao fato da sua aplicação ser limitada por barreiras operacionais, como cultura, utilidade da ferramenta de ML nesse ambiente, financiamento tecnológico e qualidade dos dados. (24) Ferramentas de ML são responsáveis por gerar alertas de gravidade e dentro de um ambiente de UTI, o excesso desses alertas, dado a gravidade dos pacientes, pode ocasionar em algo denominado fadiga de alerta, causando um aumento desproporcional da carga de trabalho dos profissionais, enquanto não fornece auxílio nas decisões clínicas. (40)

Nesse sentido, um sistema pragmático de ML que pode ser utilizado em tempo real na beira do leito, pode se mostrar uma ferramenta mais adequada para fornecer informações aos médicos intensivistas no momento da decisão de alta da UTI. Então, foi desenvolvido e validado um modelo de *machine learning point-of-care* para prever o desfecho composto de mortalidade ou readmissão na UTI em até 48 horas após a alta, utilizando informações de uma base de dados de 13 hospitais no Brasil.

1.4 MODELOS DE *MACHINE LEARNING*

Para este trabalho foram selecionados 4 modelos de *machine learning* para a avaliação do desfecho. O LGBM é bastante utilizado, dado sua fácil interpretação, alta eficiência e acurácia. (41) *TabNet* utiliza uma arquitetura de aprendizado de dados tabulares, e utiliza atenção sequencial para escolher quais características raciocinar durante as etapas de decisão, tornando o seu aprendizado mais eficiente por ser direcionado para

as características mais salientes dos dados. (42) *RuleFit* é um algoritmo que procura criar comandos simples a partir de dados complexos, de modo que a análise seja simplificada por comandos claros e fáceis de seguir. (43) O modelo FIGS foi adaptado para o contexto de decisões clínicas por levar em consideração a heterogeneidade dos dados médicos, além disso, é altamente interpretável, por utilizar uma estrutura que combina regras lógicas e adição. (44)

1.5 JUSTIFICATIVA

As taxas de readmissão na UTI nos primeiros dias após a alta e as taxas de mortalidade nas enfermarias após a alta da UTI mostram que são necessárias estratégias que possam evitar esses desfechos desfavoráveis para os pacientes que estiveram em ambiente crítico. Logo, uma ferramenta que possa ser utilizada na beira do leito pelo médico intensivista no momento de decisão da alta pode ser útil em evitar altas que poderiam gerar maior morbidade e mortalidade ao paciente e maiores custos ao sistema de saúde.

2. OBJETIVOS

2.1 OBJETIVO GERAL

Avaliar a capacidade de predição, por meio de ferramentas de *machine learning*, do desfecho composto de mortalidade e readmissão em UTI nas primeiras 48 horas após a alta.

2.2 OBJETIVOS ESPECÍFICOS

Testar o modelo *Fast Interpretable Greedy-Tree Sums* para a predição do desfecho composto.

Testar o modelo *Light Gradient Boosting Machine* para a predição do desfecho composto.

Testar o modelo *TabNet* para a predição do desfecho composto.

Testar o modelo *RuleFit* para a predição do desfecho composto.

Comparar os modelos com a capacidade de predição do escore MEWS.

Identificar quais modelos de *machine learning* apresentam maior acurácia para predição do desfecho de interesse.

3. MATERIAIS E MÉTODOS

O presente estudo foi aprovado pelo Comitê Nacional de Ética em Pesquisa (CONEP) pelo número 99706718.9.0000.0098. Os dados desse estudo foram providos pela Munai TM (Curitiba – PR, Brasil), uma empresa de tecnologia em saúde responsável por uma base de dados conectada a plataforma de prontuário eletrônico de hospitais. Para esse estudo de coorte observacional foram utilizados dados de todos os adultos que receberam cuidados em unidades de terapia intensiva em 13 hospitais entre janeiro de 2019 e julho de 2021 e concordaram com o uso de seus dados, de acordo com a Lei Geral de Proteção de Dados. Apenas internamentos com transferências para a UTI foram selecionados e com pelo menos 24 a 48 horas de seguimento.

As instituições estão localizadas em diferentes regiões geográficas do Brasil e compreendem instituições públicas e privadas.

3.1 CRITÉRIOS DE INCLUSÃO E EXCLUSÃO

Foram incluídos no estudo dados de pacientes que internaram em unidades de terapia intensiva de 13 hospitais no Brasil e que apresentaram os seguintes critérios:

Pacientes que internaram entre janeiro de 2019 e julho de 2021, pacientes que concordaram em ter seus dados utilizados e pacientes que tiveram pelo menos 24 horas de seguimento.

Foram excluídos do estudo os pacientes que apresentaram os seguintes achados:

Pacientes com menos de 18 anos, pacientes com ausência de dados de sinais vitais prévios a alta da UTI, pacientes que foram a óbito dentro do ambiente de UTI.

3.2 VARIÁVEIS DE INTERESSE

O principal desfecho de interesse foi um composto de morte ou readmissão após a alta da UTI. Readmissão foi considerada a transferência da unidade de terapia intensiva para a enfermaria e o retorno para a UTI no mesmo internamento, nas primeiras 48 horas. Duas janelas de seguimento foram observadas: as primeiras 24 horas e as primeiras 48 horas após a alta. Pacientes que permaneceram mais de 48 horas na enfermaria foram considerados como permanentes na enfermaria. A janela de 48h foi escolhida para avaliação do desfecho composto pois é considerada pelo Comitê de Indicadores de Qualidade da Sociedade de Medicina Intensiva (2) e pela sociedade europeia de medicina intensiva (45) como um indicador de segurança e qualidade.

A figura 1 demonstra a linha do tempo de observação em nosso estudo, desde a definição do desfecho, coleta de dados para a predição do desfecho e a janela de seguimento.

Para as variáveis de predição, foram extraídos dados dos prontuários médicos, informações gerais do paciente (como idade, data de admissão, data da alta, número de identificação único e motivo da alta), sinais vitais (frequência cardíaca, frequência respiratória, saturação de oxigênio, glicemia capilar e temperatura). Foram extraídas as 5 coleções de sinais vitais prévias à alta da UTI e os modelos de *machine learning* foram treinados com esses dados. A Tabela 1 demonstra as principais variáveis utilizadas pelos modelos de ML, incluindo o método principal de medição para cada variável. Combinações de recursos foram utilizadas conforme necessário, como a diferença entre valores de sinais vitais consecutivos e estatísticas a partir dos sinais vitais, como média, mediana, desvio padrão, valores máximo e mínimo.

Figura 1. Janela de tempo de dados avaliados

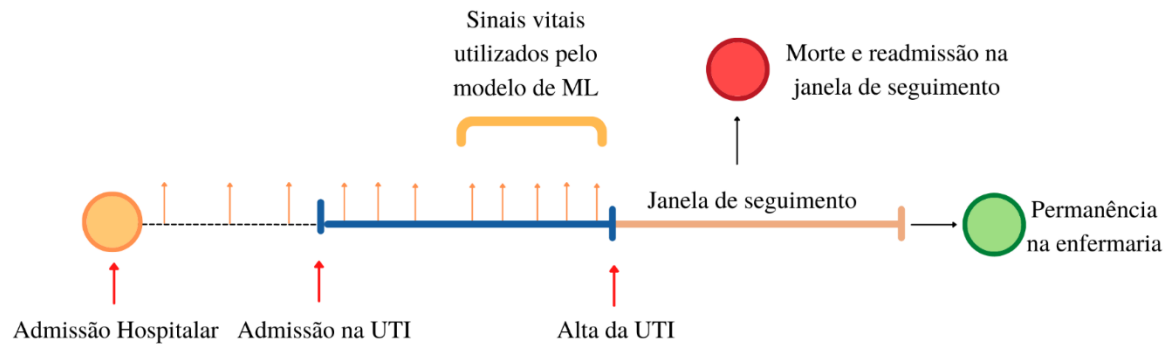


Tabela 1. Variáveis utilizadas para treinamento dos modelos de *machine learning*

Variável	Descrição
Tempo de internamento hospitalar (dias)	Espaço de tempo entre a admissão e a alta da UTI
Idade (anos)	Idade do paciente
Dias desde o último atendimento	Se o paciente já foi previamente admitido no hospital, número de dias entre as admissões
Frequência cardíaca (batimentos por minuto)	Medida da frequência cardíaca. t é o último sinal vital coletado, $t-1$ o prévio e segue essa sequência
Frequência respiratória (incursões respiratórias por minuto)	Medida da frequência respiratória. t é o último sinal vital coletado, $t-1$ o prévio e segue essa sequência
Pressão arterial diastólica (em milímetros de mercúrio)	Medida da pressão arterial diastólica. t é o último sinal vital coletado, $t-1$ o prévio e segue essa sequência
Pressão arterial sistólica (em milímetros de mercúrio)	Medida da pressão arterial sistólica. t é o último sinal vital coletado, $t-1$ o prévio e segue essa sequência

Saturação de oxigênio (%)	Medida da saturação de oxigênio. t é o último sinal vital coletado, t-1 o prévio e segue essa sequência
Glicemia capilar (miligramas por decilitro)	Medida da glicemia capilar. t é o último sinal vital coletado, t-1 o prévio e segue essa sequência
Temperatura (graus celsius)	Medida da temperatura. t é o último sinal vital coletado, t-1 o prévio e segue essa sequência
Delta de sinais vitais	Diferença entre os dados vitais coletados nos tempos t e t-1
Delta de data e hora da coleta de dados	Diferença de data e hora entre as coletas t e t-1 em segundos
Plano de saúde	1 para plano de saúde público, 0 se outro plano
Especialidade médica	Especialidade medida do atendimento. 1 para clínico e 0 para cirúrgico
Sexo	Sexo biológico do paciente. 1 para feminino e 0 para masculino

3.3 MODELOS DE MACHINE LEARNING

Para desenvolver e então comparar a performance entre os diferentes modelos de predição, o conjunto de dados foi separado entre duas coortes de forma randomizada. Os modelos foram treinados com um conjunto de 80% dos dados e 20% de dados não vistos previamente foram utilizados apenas para avaliar os modelos finais. Quatro diferentes modelos de machine learning foram utilizados: *Light Gradient Boosting Machine* (LGBM) (41), *TabNet* (42), *RuleFit* (43) e *Fast Interpretable Greedy-Tree Sums* (FIGS). (44) Cada modelo foi treinado e otimizado em uma configuração de validação cruzada aninhada, sendo realizada cinco vezes na camada externa e três vezes na parte de otimização.

Os modelos foram aplicados em dois experimentos. Um utilizando uma janela de seguimento de 24 horas e outro utilizando um período de 48 horas.

3.4 COMPARAÇÕES DE ACURÁCIA

Os modelos foram avaliados utilizando seis métricas: área sob a curva (AUC), acurácia, precisão média, escore F1, precisão e recall, por serem as métricas mais comuns utilizadas em classificação binária.

Acurácia é a razão entre as amostras corretamente classificadas e o total de amostras avaliadas. Recall também é chamado de sensibilidade ou taxa de verdadeiros positivos e denota a taxa de positivos corretamente classificados, é a razão entre as amostras corretamente classificadas como positivo e o total de amostras positivas. Precisão é a razão entre amostras classificadas corretamente e o total de amostras classificadas em uma determinada classe, para casos positivos é o valor preditivo positivo e para negativos corresponde ao valor preditivo negativo. O escore F1 é uma média harmônica de precisão e recall. (46)

AUC é utilizada como uma métrica de qualidade de um modelo, quanto maior a área sob a curva, melhor será o modelo. Essa métrica não é sensível ao desbalanço de classes. (47)

Também foi avaliada a explicabilidade de cada modelo, mostrando a interpretação mais aprofundada dos resultados. A acurácia dos nossos quatro modelos foi comparada utilizando o MEWS, como fonte de comparação com um modelo já estabelecido na literatura. (37) Para estimar o nível de consciência para o cálculo do MEWS, foi pontuado o menor valor para pacientes que tiveram o desfecho positivo e o maior valor para os pacientes que não apresentaram o desfecho.

Todas as análises foram realizadas utilizando a linguagem de programação R versão 3.3.0 (A fundação R para computação estatística). Um valor de $P < 0,05$ demonstrou significância estatística.

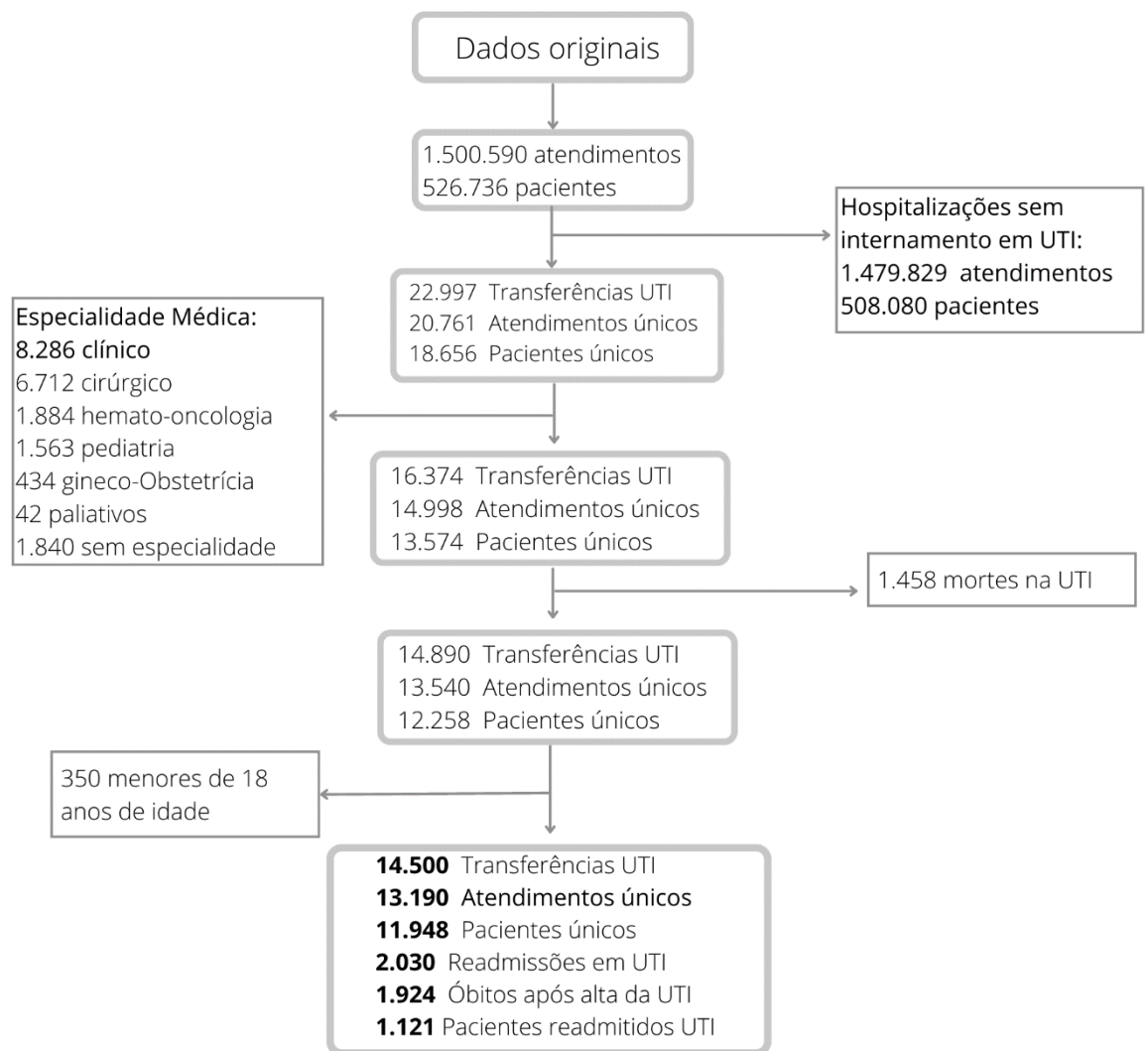
Os resultados são apresentados baseando-se nas recomendações dos Guidelines STROBE (48) e DECIDE-AI. (49)

4. RESULTADOS

4.1. PARTICIPANTES

Este estudo incluiu dados de 14.500 transferências para a UTI, sendo, desses, 13.190 atendimentos únicos, totalizando 11.948 pacientes, e tendo ocorrido 1.924 óbitos após a alta (13,2% das transferências), 2.030 readmissões na UTI (14% das transferências), e 1.121 (9,3% do total de pacientes) pacientes com registro único readmitidos na UTI. A Figura 2 demonstra o fluxograma da extração dos dados. A causa mais comum de exclusão da amostra foi a ausência de dados da admissão da UTI.

Figura 2. Fluxograma dos dados extraídos



4.2 DESCRIÇÃO DOS DADOS

A tabela 2 demonstra as características clínicas e demográficas da população do conjunto de dados final, incluindo idade, sexo, plano de saúde e especialidade médica, estratificados pelo desfecho calculado nas primeiras 48 horas após a alta da UTI (permanência na enfermaria, readmissão na unidade de terapia intensiva ou morte). Na tabela

abaixo cada atendimento único é considerado, já que idade, plano de saúde e especialidade medica podem diferir entre atendimentos para o mesmo paciente.

Foram avaliados 11.948 pacientes de 13 hospitais – públicos e privados - no Brasil com 1.924 óbitos e 2.030 readmissões na UTI em 48 horas após a alta. A tabela 3 demonstra as métricas avaliadas nas duas janelas de tempo (24 e 48 horas) para os quatro diferentes modelos de ML. Também são apresentadas as métricas para predição do desfecho utilizando o MEWS. Os dados dos pacientes excluídos do estudo podem ser visualizados na tabela 4.

Tabela 2. Características demográficas

		Permanência	Readmissão/óbito	Total	P
Idade	Média (DP)	61.5 (15.2)	62.8 (14.3)	61.8 (15.6)	0.043
Sexo (%)	Feminino	4508 (44.9)	828 (43.2)	5336 (44.6)	0.059
	Masculino	5529 (55.1)	1089 (56.8)	6618 (54.4)	
Plano de saúde (%)	Público	5219 (52)	1086 (56.6)	6305 (52.7)	0.019
	Privado	4818 (48)	831 (43.4)	5649 (47.3)	
Especialidade médica (%)	Clínica	4977 (49.6)	1305 (68)	6282 (52.5)	0.022
	Cirúrgica	5060 (50.4)	612 (32)	5672 (47.5)	
Charlson Score	Média (DP)	0.36 (0.79)	0.5 (0.7)	0.4 (0.78)	0.85
Frequência cardíaca	Média (DP)	82.1 (18)	87.9 (19.1)	83.05 (18.2)	0.002
Frequência respiratória	Média (DP)	18.6 (4.13)	20.1 (5.47)	18.8 (4.4)	0.045

Glicemia capilar	Média (DP)	154 (81.6)	159 (84.8)	154.3 (82)	0.23
Pressão arterial diastólica	Média (DP)	74.4 (14.6)	72.9 (15.7)	74.1 (14.8)	0.024
Pressão arterial sistólica	Média (DP)	127 (24)	124 (24.6)	126.6 (24)	0.015
Saturação de oxigênio	Média (DP)	95.6 (3.9)	94.3 (4.4)	95.3 (4.8)	0.009
Temperatura	Média (DP)	36.1 (0.6)	36.1 (0.77)	36 (0.7)	0.548

Tabela 3. Métricas avaliadas nos testes para os diferentes modelos de ML e MEWS

	Janela de tempo	AUC	Acurácia	Precisão média	Escore F1	Precisão	Recall
LGBM	24h	.844	.923	.418	.427	.483	.382
	48h	.824	.913	.576	.562	.666	.486
TabNet	24h	.828	.850	.428	.381	.276	.617
	48h	.817	.847	.510	.486	.395	.632
RuleFit	24h	.812	.872	.411	.402	.309	.573
	48h	.798	.779	.545	.418	.299	.691
FIGS	24h	.780	.925	.347	.352	.508	.269

	48h	.782	.902	.486	.431	.645	.324
MEWS	24h	.733	.901	.242	.346	.344	.347
	48h	.750	.891	.395	.444	.538	.378

Tabela 4. Diferenças entre os pacientes excluídos do estudo

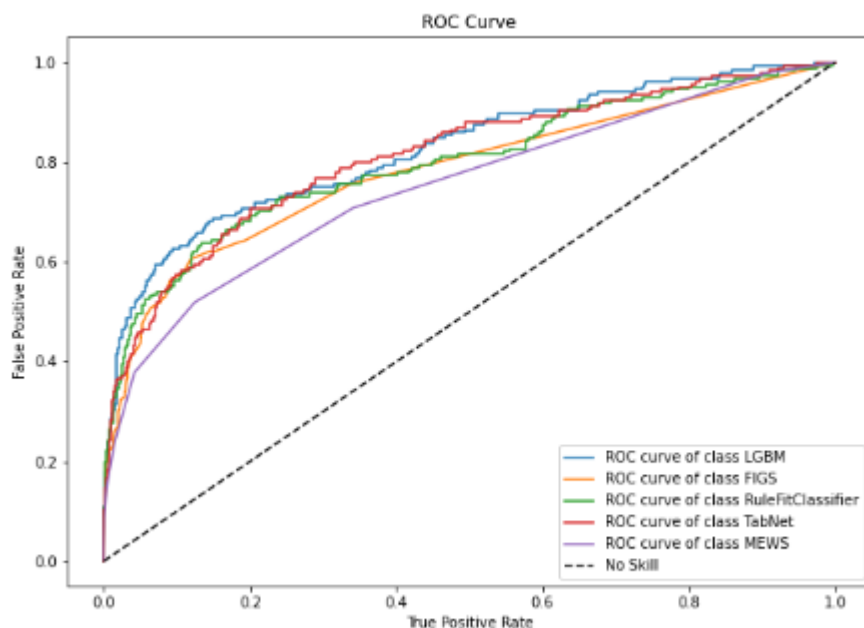
		Grupo Completo	Excluídos	P
Idade	Média (DP)	61.8 (15.8)	60.2 (15.9)	0.242
Sexo (%)	Feminino	5054 (44.5%)	282 (47.5%)	0.166
	Masculino	6306 (55.5%)	312 (52.5%)	
Plano de Saúde (%)	Privado	5220 (46%)	429 (72.2%)	0.001
	Público	6140 (54%)	165 (27.8%)	
Especialidade Médica (%)	Clínico	5887 (51.8%)	395 (66.5%)	0.001
	Cirúrgico	5473 (48.2%)	199 (33.5%)	

Como demonstrado na tabela 3, o LGBM apresentou a melhor performance em área sob a curva ROC (AUC) e escore F1 nas duas janelas de tempo estudadas, além de apresentar

a maior acurácia na janela de 48 horas. O modelo FIGS demonstrou a maior acurácia na janela de 24h. O TabNet apresentou o melhor escore de recall e precisão média na janela de tempo das primeiras 24 horas, enquanto o LGBM demonstrou a maior precisão média para a janela de 48 horas. Os modelos de ML demonstraram performance superior ao MEWS na maioria das métricas, exceto pelo TabNet e o RuleFit na métrica precisão e o FIGS em recall.

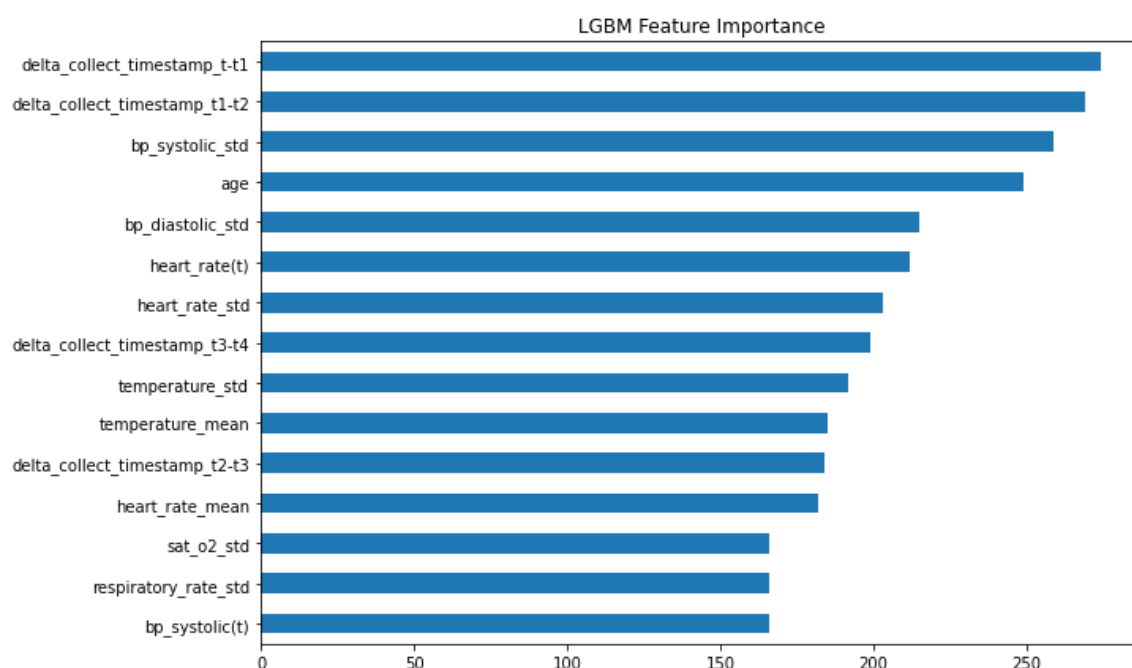
A figura 3 demonstra a curva ROC para os modelos treinados e o MEWS para prever readmissão e morte nas primeiras 48 horas após a alta da UTI. As curvas dos modelos de *machine learning* apresentaram melhor performance do que o MEWS. As melhores performances em relação a área sob a curva foram dos modelos LGBM, FIGS e TabNet.

Figura 3. Curva ROC para os diferentes modelos de *machine learning*



Entre os modelos avaliados, houve diferença de peso para as diferentes variáveis avaliadas, podendo ser um fator que contribuiu para melhor performance do modelo LGBM. Na figura 4 pode ser visualizado o peso atribuído para as variáveis por este modelo.

Figura 4. Peso atribuído para as variáveis utilizando o modelo LGBM



5. DISCUSSÃO

Neste estudo observacional multicêntrico realizado no Brasil, descobrimos que um Sistema de Suporte à Decisão Clínica (SSDC) *point-of-care* foi capaz de prever a mortalidade ou a readmissão dentro de 48 horas após a alta da UTI. Todos os modelos de *machine learning* superaram os sistemas tradicionais de alerta precoce. O modelo Light Gradient Boosting Machine (LGBM) em particular apresentou área sob a curva maior e precisão maiores do que os outros modelos de *machine learning*. Nossos resultados sugerem que um SSDC no *point-of-care* pode orientar os médicos na tomada de decisões sobre a alta, otimizando o cuidado do paciente e aumentando a segurança nas transições complexas de cuidado.

Para projetar nossos modelos, utilizamos dados das cinco últimas medições de sinais vitais registradas antes da alta da UTI. Portanto, nossos resultados representam dados próximos à decisão de alta da UTI. A transição de pacientes da UTI para a enfermaria representa uma fase de alto risco no cuidado do paciente, apresentando desafios

significativos para os médicos no leito (50). Pacientes readmitidos na UTI após a alta tendem a ter internações mais longas, taxas de mortalidade mais elevadas e custos adicionais com assistência médica (4). Portanto, há uma necessidade de estratégias aprimoradas para ajudar os médicos a decidir o momento ideal para a alta.

Escore tradicionais de alerta precoce, como NEWS, MEWS e Quick SOFA, têm sido amplamente utilizados. No entanto, as persistentes taxas elevadas de readmissão pós-alta e mortalidade destacam suas limitações (51). Nas últimas décadas, modelos de *machine learning* (ML) surgiram como auxílios potenciais nesse processo de tomada de decisão. No entanto, sua implementação em ambientes de cuidados críticos pode contribuir para a sobrecarga de pessoal e fadiga de alerta (40). Um modelo de ML *point-of-care* pode oferecer *insights* oportunos e reduzir notificações irrelevantes.

Nosso estudo se soma a um crescente corpo de literatura que descreve modelos de ML treinados com dados de sinais vitais antes da alta da UTI.

Um estudo (9) utilizou o *machine learning* (ML) para prever a readmissão, dividindo o conjunto de dados em 60% de treinamento e 40% de validação interna. Os dados foram coletados nas 24 horas que antecederam a alta da UTI, utilizando os últimos valores medidos como variáveis preditoras. Este estudo relatou uma taxa de readmissão de 11%, com o modelo de ML superando o escore MEWS nas primeiras 72 horas após a alta (AUC de 0,73 vs. 0,60).

ML também foi usado para prever readmissões (33), neste grupo, foi realizado treinamento com 80% dos dados e de validação com os outros 20%. Os dados utilizados para comparação foram aqueles obtidos no momento da admissão, durante permanência na UTI e precedendo a saída da UTI, com foco nas últimas 24 horas antes da alta da UTI. Este estudo alcançou uma AUC de 0,76 na previsão de readmissão dentro de 72 horas após a alta.

Pesquisadores relataram uma AUC de 0,70 para prever o desfecho composto de morte ou readmissão dentro de 48 horas após a alta da UTI. (18) O momento para realizar a previsão para este desfecho foi determinado como o momento da última medição de dados, normalmente ao final da estadia na UTI.

Em um estudo brasileiro (32), os desfechos de readmissão foram avaliados utilizando os dados coletados na admissão do paciente e características iniciais do internamento na UTI, resultando em uma AUC de 0,92.

Todos os modelos de *machine learning* tiveram melhor desempenho do que os modelos estáticos em outro estudo (52) quando realizada uma análise retrospectiva de dados da UTI para prever o desfecho de readmissão na UTI em 30 dias.

Da mesma forma, um estudo (53) se concentrou em readmissões dentro de uma janela de 30 dias após a alta da UTI. Eles alocaram 90% dos dados para treinamento e validação, reservando 10% para testes. Dados do histórico do paciente, incluindo medicamentos e procedimentos durante a estadia na UTI, foram utilizados, resultando em uma AUC de 0,73.

Uma AUC de 0,74 foi alcançada (54) para a predição de readmissão nas primeiras 48 horas após a alta, considerando a condição geral do paciente, sinais vitais e escores APACHE II e SOFA nas primeiras 24 horas do internamento em UTI e na transferência para a enfermaria.

Em uma revisão sistemática (55) que discute os fatores que contribuem para modelos de previsão de readmissão na UTI bem-sucedidos, estratégias-chave incluíram a incorporação de componentes temporais longitudinais, foco em populações de pacientes semelhantes e a limitação de readmissões às relacionadas diretamente com a admissão inicial.

Um estudo (56) comparou a capacidade de predição de readmissão da equipe clínica com a capacidade de predição do mesmo desfecho por um modelo de ML. O modelo utilizou os dados disponíveis no momento da alta da UTI para prever o desfecho de readmissão durante o mesmo internamento. AUC da intuição clínica foi de 0,71 , enquanto a AUC do modelo de ML foi de 0,78 e a AUC do ML combinado com intuição clínica foi de 0,79, sendo encontrada uma diferença significativa entre a intuição clínica isolada e o uso de ML ou o uso de ML combinado com a intuição clínica.

Diversos desenhos de estudo visam avaliar os resultados após a alta da UTI, incluindo mortalidade e readmissão. Esses estudos diferem em sua metodologia, principalmente no foco do momento ideal de análise dados, que pode ser na admissão à UTI, ao longo da permanência na UTI ou imediatamente antes da alta. Também existe uma heterogeneidade no período de seguimento após a alta. Essas diferenças podem contribuir para os resultados variados observados na pesquisa.

Como resultado, muitos estudos sugerem que os modelos de *machine learning* podem superar escores clínicos tradicionais como o MEWS. Prever resultados após a alta da UTI é complexo e uma ferramenta com alta precisão para avaliar o risco de morte ou readmissão após a alta pode auxiliar significativamente os médicos em sua tomada de decisão.

Este presente estudo tem várias vantagens notáveis. Em primeiro lugar, foi utilizado um banco de dados abrangente e diversificado que representa tanto configurações de saúde privadas quanto públicas em todo o Brasil, sugerindo ampla aplicabilidade das descobertas. Além disso, a granularidade e a natureza contínua dos dados sobre parâmetros de sinais vitais acrescentam profundidade e especificidade à nossa análise. Importante, os modelos foram

treinados com dados próximos à alta da UTI, alinhando a saída do modelo com o momento crítico da tomada de decisão.

Em relação às limitações, este estudo foi realizado com dados obtidos através de uma fonte de dados secundária, o que caracteriza uma limitação importante, visto que isso é dependente da qualidade dos dados originais. Do total da amostra, 594 (4,7%) pacientes foram excluídos por falta de dados da admissão na UTI, tornando possível que uma das limitações desse estudo seja um viés de seleção da amostra. A média de readmissões de 1,8 por paciente demonstra que existiram pacientes que tiveram mais de uma readmissão na UTI, e isso também está relacionado com pacientes que apresentaram readmissões na UTI em internamentos diferentes, o que leva a um questionamento sobre a gravidade e o estado basal desses pacientes, que não foi avaliado pelos modelos de ML.

Os dados também foram obtidos de maneira retrospectiva, o que faz com que sejam necessárias novas avaliações desses modelos de maneira prospectiva para validação dos modelos para uso dentro do ambiente de UTI. Em relação aos hospitais, não foi possível avaliar se os hospitais já possuem um protocolo de predição de readmissões no momento da alta e os diferentes perfis dos pacientes das instituições representadas nesse estudo pode não ser representativo de outras populações, fazendo com que esses resultados não possam ser generalizados.

Na literatura atual existem estudos com modelos de inteligência artificial e *machine learning* que avaliaram a capacidade de predição de morte ou readmissão em UTI. Porém, o desenho desses estudos, principalmente em relação ao momento ideal de avaliação do desfecho ainda é uma variável que precisa ser melhor estudada.

6. CONCLUSÃO

Neste estudo observacional multicêntrico realizado no Brasil, descrevemos um sistema de suporte de decisão clínica *point-of-care* com capacidade de predição do desfecho de readmissão na UTI ou a morte em até 48h da alta com até 92% de precisão.

Os modelos de ML consistentemente superaram os sistemas tradicionais de alerta precoce e os resultados foram robustos tanto nas primeiras 24 quanto nas 48 horas após a alta da UTI. O modelo LGBM demonstrou a maior área sobre a curva ROC. Ensaios clínicos randomizados para avaliar a eficácia e a segurança da implementação de um sistema de suporte de decisão clínica *point-of-care* na transição de cuidados do ambiente de UTI para as enfermarias se mostram necessários.

7. REFERÊNCIAS

1. Center for Medicare and Medicaid Services. Medicare Program; Changes to the Hospital Inpatient Prospective Payment Systems and Fiscal Year 2009 Rates; Payments for Graduate Medical Education in Certain Emergency Situations; Changes to Disclosure of Physician Ownership in Hospitals and Physician Self-Referral Rules; Updates to the Long-Term Care Prospective Payment System; Updates to Certain IPPS-Excluded Hospitals; and Collection of Information Regarding Financial Relationships Between Hospitals. Final rules. Fed Regist. 2008; 73(161): 48433–9084. Available: <http://www.cms.gov/AcuteInpatientPPS/downloads/CMS-1390-F.pdf>. PMID: 18956499
2. Angus DC. Grappling with intensive care unit quality-does the readmission rate tell us anything? Crit Care Med. 1998 Nov;26(11):1779-80. doi: 10.1097/00003246-199811000-00008. PMID: 9824061.
3. Rojas JC, Lyons PG, Jiang T, Kilaru M, McCauley L, Picart J, Carey KA, Edelson DP, Arora VM, Churpek MM. Accuracy of Clinicians' Ability to Predict the Need for Intensive Care Unit Readmission. Ann Am Thorac Soc. 2020 Jul;17(7):847-853. doi: 10.1513/AnnalsATS.201911-828OC. PMID: 32125877; PMCID: PMC7328179.
4. Nates JL, Nunnally M, Kleinpell R, et al. ICU Admission, Discharge, and Triage Guidelines: A Framework to Enhance Clinical Operations, Development of Institutional Policies, and Further Research. Critical care medicine. 2016. V 44 N 8. Pag 1553-1602.
5. Nishi GK, Suh RH, Wilson MT, Cunneen SA, Margulies DR, Shabot MM. Analysis of causes and prevention of early readmission to surgical intensive care. Am Surg. 2003 Oct;69(10):913-7. PMID: 14570374.
6. Levy MF, Greene L, Ramsay MA, Jennings LW, Ramsay KJ, Meng J, Hein HA, Goldstein RM, Husberg BS, Gonwa TA, Klintmalm GB. Readmission to the intensive care unit after liver transplantation. Crit Care Med. 2001 Jan;29(1):18-24. doi: 10.1097/00003246-200101000-00004. PMID: 11176152.

7. Elliott M. Readmission to intensive care: a review of the literature. *Aust Crit Care*. 2006 Aug;19(3):96-8, 100-4. doi: 10.1016/s1036-7314(06)80004-4. PMID: 16933636.
8. Álvarez EET, Mediavilla KAG, Solís C R, González NV, Balanza JAL. Factores de riesgo de reingreso en UCI y análisis de la mortalidad intrahospitalaria. *Medicina Clínica (Barc)*. 2021.
9. Rojas JC, Carey KA, Edelson DP, Venable LR, et al. Predicting Intensive Care Unit Readmission with Machine Learning Using Electronic Health Record Data. *Annals of the American Thoracic Society*. 2018. V15, N 7. Pag 846-853.
10. Morgan R, Lloyd-Williams F, Wright M, Morgan-Warren RJ. An early warning scoring system for detecting developing critical illness. 1997. <https://www.scienceopen.com/document?vid=28251d22-8476-40a6-916d-1a34796816e4>
11. Gerry S, Bonnici T, Birks J, Kirtley S, Virdee PS, Watkinson PJ, Collins GS. Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. *BMJ*. 2020 May 20;369:m1501. doi: 10.1136/bmj.m1501. PMID: 32434791; PMCID: PMC7238890.
12. Stenhouse C, Coates S, Tivey M, Allsop P, Parker T. Prospective evaluation of a modified Early Warning Score to aid earlier detection of patients developing critical illness on a general surgical ward. *Br J Anaesth*. 1999;84:663P.
13. Vergara P, Forero D, Bastidas A, Garcia JC, Blanco J, Azocar J, Bustos RH, Liebisch H. Validation of the National Early Warning Score (NEWS)-2 for adults in the emergency department in a tertiary-level clinic in Colombia: Cohort study. *Medicine (Baltimore)*. 2021 Oct 8;100(40):e27325. doi: 10.1097/MD.00000000000027325. PMID: 34622831; PMCID: PMC8500632.
14. Churpek MM, Snyder A, Han X, Sokol S, Pettit N, Howell MD, Edelson DP. Quick Sepsis-related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores for Detecting Clinical Deterioration in Infected Patients outside the Intensive Care Unit. *Am J Respir Crit Care Med*. 2017 Apr 1;195(7):906-911. doi: 10.1164/rccm.201604-0854OC. PMID: 27649072; PMCID: PMC5387705.)

15. Khwannimit B, Bhurayanontachai R, Vattanavanit V. Comparison of the accuracy of three early warning scores with SOFA score for predicting mortality in adult sepsis and septic shock patients admitted to intensive care unit. *Heart Lung*. 2019 May-Jun;48(3):240-244. doi: 10.1016/j.hrtlng.2019.02.005. Epub 2019 Mar 19. PMID: 30902348
16. Larvin M, McMahon MJ. APACHE-II score for assessment and monitoring of acute pancreatitis. *Lancet*. 1989 Jul 22;2(8656):201-5. doi: 10.1016/s0140-6736(89)90381-4. PMID: 2568529.)
17. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA*. 2001;286(14):1754-1758. doi: 10.1001/jama.286.14.1754
18. Desautels T, Das R, Calvert J, et al. Prediction of early unplanned intensive care unit readmission in a UK tertiary care hospital: a cross-sectional machine learning approach *BMJ Open* 2017;7:e017199. doi: 10.1136/bmjopen-2017-017199.
19. Durbin CG Jr, Kopel RF. A case-control study of patients readmitted to the intensive care unit. *Crit Care Med* 1993;21:1547–1553.6
20. Horrigan TP. Readmission of patients to the surgical intensive care unit: patient profiles and possibilities for prevention. *Crit Care Med*. 1985;13:961–964.7
21. Al-Jaghbeer MJ, Tekwani SS, Gunn SR, Kahn JM. Incidence and etiology of potentially preventable ICU Readmissions. *Crit Care Med*. 2016;44:1704–1709
22. Yu, KH., Beam, A.L. & Kohane, I.S. Artificial intelligence in healthcare. *Nat Biomed Eng* 2, 719–731 (2018). <https://doi.org/10.1038/s41551-018-0305-z>
23. Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism*. 2017 Apr;69S:S36-S40. doi: 10.1016/j.metabol.2017.01.011. Epub 2017 Jan 11. PMID: 28126242.
24. Y. Raita, T. Goto, M.K. Faridi, D.F.M. Brown, C.A. Camargo Jr, K. Hasegawa. Emergency department triage prediction of clinical outcomes using machine learning models. *Crit Care*, 23 (2019), p. 64

25. W. Hong, A. Haimovich, A.R. Taylor. Predicting hospital admission at emergency department triage using machine learning. *PLoS One*, 13 (2018), p. e0201016.
26. S. Levin, M. Toerper, E. Hamrock, et al. Machine-learning-based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the Emergency Severity Index. *Ann Emerg Med*, 71 (2017), pp. 565-574.
27. H.M. Giannini, J.C. Ginestra, C. Chivers, et al. A machine learning algorithm to predict severe sepsis and septic shock: development, implementation, and impact on clinical practice. *Crit Care Med*, 47 (2019), pp. 1485-1492.
28. J. Kobylarz Ribeiro et al., "A Machine Learning Early Warning System: Multicenter Validation in Brazilian Hospitals," 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS), Rochester, MN, USA, 2020, pp. 321-326, doi: 10.1109/CBMS49503.2020.00067)
29. Adlung L, Cohen Y, Mor U, Elinav E. Machine learning in clinical decision making. *Med (N Y)*. 2021 Jun 11;2(6):642-665. doi: 10.1016/j.medj.2021.04.006. Epub 2021 Apr 30. PMID: 35590138.
30. Arfat Y, Mittone G, Esposito R, Cantalupo B, DE Ferrari GM, Aldinucci M. Machine learning for cardiology. *Minerva Cardiol Angiol*. 2022 Feb;70(1):75-91. doi: 10.23736/S2724-5683.21.05709-4. Epub 2021 Aug 2. PMID: 34338485.
31. Peiffer-Smadja N, Rawson TM, Ahmad R, Buchard A, Georgiou P, Lescure FX, Birgand G, Holmes AH. Machine learning for clinical decision support in infectious diseases: a narrative review of current applications. *Clin Microbiol Infect*. 2020 May;26(5):584-595. doi: 10.1016/j.cmi.2019.09.009. Epub 2019 Sep 17. Erratum in: *Clin Microbiol Infect*. 2020 Aug;26(8):1118. PMID: 31539636.
32. Loreto M, Lisboa T, Moreira VP. Early prediction of ICU readmissions using classification algorithms. *Comput Biol Med*. 2020 Mar;118:103636. doi: 10.1016/j.combiomed.2020.103636. Epub 2020 Feb 1. PMID: 32174313.
33. Pakbin A, Rafi P, Hurley N, Schulz W, Harlan Krumholz M, Bobak Mortazavi J. Prediction of ICU Readmissions Using Data at Patient Discharge. *Annu Int Conf IEEE Eng*

Med Biol Soc. 2018 Jul;2018:4932-4935. doi: 10.1109/EMBC.2018.8513181. PMID: 30441449.

34. A.S. Fialho, F. Cismondi, S.M. Vieira, S.R. Reti, J.M.C. Sousa, S.N. Finkelstein, Data mining using clinical physiology at discharge to predict ICU readmissions, *Expert Systems with Applications*, Volume 39, Issue 18, 2012, Pages 13158-13165, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2012.05.086>.

35. Lior Turgeman, Jerrold H. May, A mixed-ensemble model for hospital readmission, *Artificial Intelligence in Medicine*, Volume 72, 2016, Pages 72-82, ISSN 0933-3657, <https://doi.org/10.1016/j.artmed.2016.08.005>.

36. Thorat PJ, Fornasa M, de Bruin DP, Tonutti M, Hovenkamp H, Driessen RH, Girbes ARJ, Hoogendoorn M, Elbers PWG. Explainable Machine Learning on Amsterdam IMCdb for ICU Discharge Decision Support: Uniting Intensivists and Data Scientists. *Crit Care Explor*. 2021 Sep 10;3(9):e0529. doi: 10.1097/CCE.0000000000000529. PMID: 34589713; PMCID: PMC8437217.

37. McWilliams CJ, Lawson DJ, Santos-Rodriguez R, et al. Towards a decision support tool for intensive care discharge: machine learning algorithm development using electronic healthcare data from MIMIC-III and Bristol, UK. *BMJ Open* 2019;9:e025925. doi: 10.1136/bmjopen-2018-025925.

38. Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. *J Intern Med*. 2018 Dec;284(6):603-619. doi: 10.1111/joim.12822. Epub 2018 Sep 3. PMID: 30102808.

39. Watson J, Huttyra CA, Clancy SM, Chandiramani A, Bedoya A, Ilangovan K, Nderitu N, Poon EG. Overcoming barriers to the adoption and implementation of predictive modeling and machine learning in clinical care: what can we learn from US academic medical centers? *JAMIA Open*. 2020 Apr 10;3(2):167-172. doi: 10.1093/jamiaopen/ooz046. PMID: 32734155; PMCID: PMC7382631.

40. Poly TN, Islam MM, Muhtar MS, Yang HC, Nguyen PAA, Li YJ. Machine Learning Approach to Reduce Alert Fatigue Using a Disease Medication-Related Clinical Decision Support System: Model Development and Validation. *JMIR Med Inform*. 2020 Nov 19;8(11):e19489. doi: 10.2196/19489. PMID: 33211018; PMCID: PMC7714650.

41. Guolin K, Qi M, Thomas F, Taifeng W, Wei C, Weidong M, Qiwei Y, Tie-Yan L. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Advances in Neural Information Processing Systems* 30 (NIPS 2017). ISBN: 9781510860964
42. Arik, S. Ö., & Pfister, T. (2021). TabNet: Attentive Interpretable Tabular Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(8), 6679-6687. <https://doi.org/10.1609/aaai.v35i8.16826>
43. Jerome H. Friedman, Bogdan E. Popescu "Predictive learning via rule ensembles," *The Annals of Applied Statistics*, Ann. Appl. Stat. 2(3), 916-954, (September 2008)
44. Tan, Yan Shuo, Singh, Chandan, Nasser, Keyan, Agarwal, Abhineet, Duncan, James, Ronen, Omer, Epland, Matthew, Kornblith, Aaron, and Yu, Bin. Fast Interpretable Greedy-Tree Sums (FIGS). 2022. Retrieved from <https://par.nsf.gov/biblio/10343646>. ArXiv.org
45. Rhodes A, Moreno RP, Azoulay E, Capuzzo M, Chiche JD, Eddleston J, Endacott R, Ferdinande P, Flaatten H, Guidet B, Kühlen R, León-Gil C, Martín Delgado MC, Metnitz PG, Soares M, Sprung CL, Timsit JF, Valentin A; Task Force on Safety and Quality of European Society of Intensive Care Medicine (ESICM). Prospectively defined indicators to improve the safety and quality of care for critically ill patients: a report from the Task Force on Safety and Quality of the European Society of Intensive Care Medicine (ESICM). *Intensive Care Med*. 2012 Apr;38(4):598-605. doi: 10.1007/s00134-011-2462-3. Epub 2012 Jan 26. PMID: 22278594.
46. Hicks SA, Strümke I, Thambawita V, Hammou M, Riegler MA, Halvorsen P, Parasa S. On evaluation metrics for medical applications of artificial intelligence. *Sci Rep*. 2022 Apr 8;12(1):5979. doi: 10.1038/s41598-022-09954-8. PMID: 35395867; PMCID: PMC8993826.
47. Ting KM. *Encyclopedia of machine learning*. Springer. 2011. ISBN 978-0-387-30164-8.
48. Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, Shilton A, Yearwood J, Dimitrova N, Ho TB, Venkatesh S, Berk M. Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View. *J Med Internet Res*. 2016 Dec 16;18(12):e323. doi: 10.2196/jmir.5870. PMID: 27986644; PMCID: PMC5238707.

49. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, Denniston AK, Faes L, Geerts B, Ibrahim M, Liu X, Mateen BA, Mathur P, McCradden MD, Morgan L, Ordish J, Rogers C, Saria S, Ting DSW, Watkinson P, Weber W, Wheatstone P, McCulloch P; DECIDE-AI expert group. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *BMJ*. 2022 May 18;377:e070904. doi: 10.1136/bmj-2022-070904. PMID: 35584845; PMCID: PMC9116198.)
50. Kripalani S, Jackson AT, Schnipper JL, Coleman EA. Promoting effective transitions of care at hospital discharge: a review of key issues for hospitalists. *J Hosp Med*. 2007 Sep;2(5):314-23. doi: 10.1002/jhm.228. PMID: 17935242.
51. Vollam S, Dutton S, Lamb S, Petrinic T, Young JD, Watkinson P. Out-of-hours discharge from intensive care, in-hospital mortality and intensive care readmission rates: a systematic review and meta-analysis. *Intensive Care Med*. 2018 Jul;44(7):1115-1129. doi: 10.1007/s00134-018-5245-2. Epub 2018 Jun 25. PMID: 29938369; PMCID: PMC6061448.
52. Venugopalan J, Chanani N, Maher K, Wang MD. Combination of static and temporal data analysis to predict mortality and readmission in the intensive care. *Annu Int Conf IEEE Eng Med Biol Soc*. 2017 Jul;2017:2570-2573. doi: 10.1109/EMBC.2017.8037382. PMID: 29060424; PMCID: PMC7370856.
53. Barbieri S, Kemp J, Perez-Concha O, Kotwal S, Gallagher M, Ritchie A, Jorm L. Benchmarking Deep Learning Architectures for Predicting Readmission to the ICU and Describing Patients-at-Risk. *Sci Rep*. 2020 Jan 24;10(1):1111. doi: 10.1038/s41598-020-58053-z. PMID: 31980704; PMCID: PMC6981230.
54. Yin YL, Sun MR, Zhang K, Chen YH, Zhang J, Zhang SK, Zhou LL, Wu YS, Gao P, Shen KK, Hu ZJ. Status and Risk Factors in Patients Requiring Unplanned Intensive Care Unit Readmission Within 48 Hours: A Retrospective Propensity-Matched Study in China. *Risk Manag Healthc Policy*. 2023 Mar 11;16:383-391. doi: 10.2147/RMHP.S399829. PMID: 36936882; PMCID: PMC10015949.
55. Ruppert MM, Loftus TJ, Small C, Li H, Ozrazgat-Baslanti T, Balch J, Holmes R, Tighe PJ, Upchurch GR Jr, Efron PA, Rashidi P, Bihorac A. Predictive Modeling for Readmission to Intensive Care: A Systematic Review. *Crit Care Explor*. 2023 Jan 6;5(1):e0848. doi: 10.1097/CCE.0000000000000848. PMID: 36699252; PMCID: PMC9829260.

56. Rojas, JC et al. Man vs Machine: Comparison of a Machine Learning Algorithm to Clinician Intuition for Predicting Intensive Care Unit Readmission. American Thoracic Society. 2019. DOI:10.1164/ajrccm-conference.2019.199.1_MeetingAbstracts.A2459