Development and Evaluation of an Automated Machine Learning Algorithm for In-Hospital Mortality Risk Adjustment Among Critical Care Patients\* Ryan J. Delahanty, PhD1 ; David Kaufman, MD, FCCM2 ; Spencer S. Jones, PhD1

Objetivos: Algoritmos de ajuste de risco para mortalidade na UTI são necessários para medir e melhorar o desempenho da UTI. Os algoritmos de ajuste de risco existentes não são amplamente adotados. As principais barreiras à adoção incluem licenciamento e custos de implementação, bem como custos de mão de obra associados à coleta de dados com uso intensivo de pessoas. A adoção generalizada de registros eletrônicos de saúde torna o ajuste automatizado de risco viável. Usando métodos modernos de aprendizado de máquina e ferramentas de código aberto, desenvolvemos e avaliamos um algoritmo de ajuste de risco retrospectivo para mortalidade intra-hospitalar entre pacientes de UTI. O escore de Risco de Morte em Paciente Interno pode ser totalmente automatizado e depende de elementos de dados que são gerados no curso de processos hospitalares usuais.

Local: Cento e trinta e uma UTIs em 53 hospitais operados pela Tenet Healthcare.

Pacientes: Uma coorte de 237.173 pacientes da UTI com alta entre janeiro de 2014 e dezembro de 2016.

Desenho: Os dados foram divididos aleatoriamente em conjuntos de dados de treinamento (36 hospitais) e validação (17 hospitais). A seleção de recursos e o treinamento do modelo foram realizados usando o conjunto de treinamento enquanto a discriminação, calibração e precisão do modelo foram avaliadas no conjunto de dados de validação.

Medidas e resultados principais: A discriminação do modelo foi avaliada com base na área sob a curva característica de operação do receptor; a precisão e a calibração foram avaliadas por meio de escores de Brier ajustados e análise visual das curvas de calibração. Dezessete recursos, incluindo uma mistura de elementos de dados clínicos e administrativos, foram mantidos no modelo final. O escore de risco de morte hospitalar demonstrou excelente discriminação (área sob a curva característica de operação do receptor = 0,94) e calibração (escore de Brier ajustado = 52,8%) no conjunto de dados de validação; esses resultados se comparam favoravelmente às estatísticas de desempenho publicadas para os algoritmos de ajuste de risco de mortalidade mais comumente usados.

Conclusões: A baixa adoção de algoritmos de ajuste de risco de mortalidade em UTI impede o progresso no sentido de aumentar o valor dos cuidados de saúde prestados em UTI. O escore de Risco de Morte em Paciente Interno tem muitos atributos atraentes que abordam as principais barreiras para a adoção de algoritmos de ajuste de risco na UTI e tem um desempenho comparável aos algoritmos de uso intensivo de humanos existentes. Os algoritmos de ajuste de risco automatizado têm o potencial de evitar barreiras conhecidas à adoção, como taxas de licenciamento proibitivas e custos diretos de mão de obra significativos. Avaliações adicionais são necessárias para garantir que o nível de desempenho observado neste estudo possa ser alcançado em locais independentes. (Crit Care Med 2018; 46: e481-e488)

**Desenvolvimento e avaliação de modelo**

O objetivo do modelo era estimar de forma confiável o risco de morte hospitalar em pacientes de UTI. Para este fim, XGBoost (v. 0.4; T. Chen, T. He, M. Benesty; https://CRAN.R-project.org/package=xgboost), uma biblioteca de software de código aberto disponível gratuitamente para máquina o aprendizado foi usado para construir um modelo que pudesse estimar com segurança o risco de mortalidade de pacientes internados (20, 21). O modelo foi treinado com pacientes de 36 dos 53 hospitais (146.982 pacientes). Os pacientes dos 17 hospitais restantes (90.191 pacientes) foram “retidos” para servir como um conjunto de validação. A atribuição ao conjunto de treinamento versus validação foi determinada por meio de um estudo de simulação em que geramos 200 divisões aleatórias da população do hospital. Em seguida, calculamos as taxas de mortalidade observadas / esperadas para os conjuntos de treinamento e validação para cada divisão aleatória; em seguida, selecionamos a divisão que minimizou a diferença na mortalidade observada / esperada entre os conjuntos de treinamento e validação. O objetivo deste exercício de simulação foi reduzir o risco de viés que poderia ocorrer se uma parcela desproporcional de UTIs de baixa ou alta mortalidade fosse atribuída ao conjunto de treinamento ou validação.

**LIMITAÇÕES**

Conforme observado anteriormente, os elementos de dados administrativos, como códigos APRDRG e seus escores de risco de mortalidade e gravidade associados, são componentes importantes do escore RIPD. A precisão da codificação administrativa depende da experiência do codificador; entretanto, os regulamentos de credenciamento e codificação rigorosos limitam a variação na prática e na precisão da codificação clínica, e há um forte precedente para o uso de códigos APRDRG para ajuste de risco (17, 29). Acreditamos que a maioria dos hospitais usa códigos APR-DRG; entretanto, hospitais que não usam códigos APR-DRG não poderão usar o algoritmo RIPD. Para resolver essa limitação, também desenvolvemos e validamos uma versão do algoritmo RIPD que não incorpora códigos APR-DRG ou MS-DRG (RIPD\_reduzido). O algoritmo RIPD\_reduzido também fornece uma discriminação muito boa (AUC = 0,91) e calibração (Brier Score ajustado =

0,49). Mais detalhes relacionados ao algoritmo RIPD\_reduced podem ser encontrados nas Tabelas 3 e 4 do Apêndice (Conteúdo Digital Suplementar 1, <http://links.lww.com/CCM/D222>).

(...)continua

**CONCLUSÕES**

O estudo demonstra a viabilidade de desenvolver um algoritmo de ajuste de risco automatizado em UTI (ICU) usando métodos machine learning usando apenas elementos de dados administrativos e EHR altamente disponíveis. Mostramos que esse algoritmo, quando aplicado em um grande conjunto de UTIs de hospitais, clinicamente e geograficamente diverso, fornece excelente discriminação e calibração. A discriminação e calibração da pontuação RIPD se compara favoravelmente com os algoritmos de ajuste de risco de UTI mais estabelecidos e oferece recursos atraentes adicionais que a distinguem das ferramentas existentes, incluindo nenhum custo de mão de obra direto associado à coleta manual de dados, sem taxas de licenciamento e confiança em dados amplamente disponíveis elementos A maior parte do foco na literatura tem sido nos detalhes técnicos de tentar obter o modelo “certo” (10); no entanto, o fato de a grande maioria dos hospitais não usar nenhum ajuste de risco na UTI indica que a remoção das barreiras de tempo e custo deve ser uma área de foco mais proeminente (2). A economia potencial de custo e tempo proporcionada pelo uso de um modelo de ajuste de risco automatizado com base em software de código aberto disponível gratuitamente deve ser encorajador para todos os interessados ​​em aumentar a transparência e medição do desempenho da UTI.