**Development and Validation of an Electronic Health Record–Based Machine Learning Model to Estimate Delirium Risk in Newly Hospitalized Patients Without Known Cognitive Impairment**

IMPORTÂNCIA: Os métodos atuais de identificação de pacientes hospitalizados com risco acrescido de delírio requerem questionários administrados por enfermeiras com precisão moderada.

OBJECTIVO: Desenvolver e validar um modelo de aprendizagem automática que preveja o risco de delírio incidente com base em dados de saúde electrónicos disponíveis no momento da admissão.

DESIGN, CONFIGURAÇÃO E PARTICIPANTES Estudo de coorte retrospectivo avaliando 5 aprendizagens automáticas algoritmos para prever o delírio usando 796 variáveis clínicas identificadas por um painel de peritos como relevantes para a previsão de delírios e consistentemente disponível em registos de saúde electrónicos no prazo de 24 horas após admissão. O conjunto de formação incluiu 14 227 pacientes adultos com unidade de cuidados não intensivos hospitalar e nenhum delírio na admissão que foram dispensados entre 1 de Janeiro de 2016, e 31 de Agosto, 2017, da UCSF Health, uma grande instituição de saúde académica. O conjunto de testes incluiu 3996 pacientes com internamento hospitalar que tiveram alta entre 1 de Agosto de 2017, e 30 de Novembro de 2017.

EXPOSIÇÕES Características demográficas dos doentes, diagnósticos, registos de enfermagem, resultados laboratoriais, e medicamentos disponíveis nos registos de saúde electrónicos durante a hospitalização.

PRINCIPAIS RESULTADOS E MEDIDAS O Delirium foi definido como uma triagem positiva de enfermagem Escala ou Método de Avaliação de Confusão para a pontuação da Unidade de Cuidados Intensivos. Os modelos foram avaliados utilizando a área sob a curva característica de funcionamento do receptor (AUC) e comparada com a Sistema de pontuação de 4 pontos AWOL (idade >79 anos, falha em soletrar o mundo para trás, desorientação para colocar, e maior gravidade da doença), uma ferramenta validada de avaliação de risco de delírio rotineiramente administrados nesta coorte.

RESULTADOS O conjunto de formação incluiu 14 227 pacientes (5113 [35,9%] com idade >64 anos; 7335 [51,6%] feminino; 687 [4,8%] com delírio), e o conjunto de teste incluía 3996 pacientes (1491 [37,3%] com idade >64 anos; 1966 [49,2%] mulher; 191 [4,8%] com delírio). No total, a análise incluiu 18 223 hospitais admissões (6604 [36,2%] com idade >64 anos; 9301 [51,0%] do sexo feminino; 878 [4,8%] com delírio). O O sistema AWOL atingiu uma AUC de base de 0,678. O modelo da máquina de reforço de gradiente realizado melhor, com um AUC de 0,855. Fixando a especificidade em 90%, o modelo tinha um 59,7%(95%CI, 52,4%-66,7%) sensibilidade, 23,1%(95%CI, 20,5%-25,9%) valor preditivo positivo, 97,8%(95%CI, 97,4%-98,1%) valor preditivo negativo, e um número necessário para rastrear de 4,8. Logística penalizada regressão e modelos florestais aleatórios também tiveram bons resultados, com CUA de 0,854 e 0,848, respectivamente.

CONCLUSÕES E RELEVÂNCIA A aprendizagem mecânica pode ser utilizada para estimar o risco de delírio adquirido no hospital, utilizando dados de registos de saúde electrónicos disponíveis no prazo de 24 horas após a admissão no hospital. Tal modelo pode permitir uma orientação mais precisa dos recursos de prevenção do delírio sem aumentar a carga sobre os profissionais de saúde.

**Introdução:**

As estratégias de prevenção, no entanto, não são farmacológicas e, por conseguinte, são intensivas em recursos e pessoal.4 Uma previsão precisa do risco de delírio poderia permitir uma orientação mais precisa dos doentes de alto risco e, assim, uma maior gestão de recursos e, potencialmente, melhores resultados para os doentes.

Por exemplo, UCSF Health (a Universidade da Califórnia), São Francisco, sistema do Centro Médico) utiliza a ferramenta de rastreio AWOL para calcular o risco de delírio para pacientes recém admitidos.12 Este instrumento atribui 1 ponto para cada um dos seguintes critérios: idade superior a 79 anos; incapacidade de soletrar o mundo ao contrário; desorientação para a cidade, estado, condado, nome de hospital, ou piso; e a gravidade da doença é moderada ou grave. Uma pontuação de 2 pontos ou superior indica elevada risco e ajuda a dirigir os recursos hospitalares para a prevenção de delírios (por exemplo, serviços de reabilitação, doentes assistentes de cuidados, voluntários). Um recente estudo de coorte prospectivo na nossa instituição encontrou AWOL atingiu um AUC de 0,73 em pacientes hospitalizados com 50 anos ou mais.13

Além disso, os instrumentos existentes recapitulam bem os fatores de risco de delírio estudados, tais como a deficiência cognitiva na linha de base, o delírio na admissão, e doenças graves. 5-13

**Formação e Validação de Modelos**

Testamos o desempenho de 5 modelos de aprendizagem de máquinas em comparação com a AWOL. Os algoritmos (implementação do pacote R) incluíam regressão logística penalizada (glmnet), máquina de reforço de gradiente (gbm), rede neural artificial com uma única camada oculta (nnet), máquina vectorial de suporte linear (e1071), e floresta aleatória (randomForest). Utilizando o pacote de caretes R,19 hiperparâmetros para cada modelo foram optimizados com 3 repetições de validação cruzada de 5 vezes, cabendo depois a todo o conjunto de treino. Avaliamos então cada modelo através do cálculo da AUC no conjunto de testes completo e no subconjunto de hospitalizações em que foi realizado um AWOL. O modelo de relatório está em conformidade com a diretiva Transparente de Relatório de um Modelo de Previsão Multivariável para Prognóstico ou Diagnóstico Individual (TRIPOD).20 O código e os modelos foram disponibilizados em <https://github.com/ayoung01/delirium>.

**Discussão**

Este estudo demonstra que os modelos de aprendizagem de máquinas superam as ferramentas clínicas atuais utilizadas para avaliar o risco de delírio. Em comparação com AWOL, que se verificou ter um NNS de 11,1 no limiar de AWOL maior ou igual a 2, o nosso modelo de GBM atinge um NNS de 4,8 mantendo uma sensibilidade mais elevada do que AWOL, sugerindo que menos de metade dos pacientes precisariam de ser tratados por 1 para beneficiar de intervenções de prevenção do delírio. Modelos de aprendizagem de máquinas têm a vantagem adicional de não necessitarem de um profissional de saúde para realizar uma cabeceira avaliação do risco de delírio.

Embora o delírio seja geralmente considerado como afetando desproporcionadamente os idosos, também ocorre em doentes mais jovens, com uma prevalência de 4,7%22 e uma incidência tão elevada como 14% em grupos de alto risco.23 Ao contrário de estudos anteriores que se concentram apenas em populações mais idosas, o nosso estudo não exclui doentes com base na idade. No limiar de 90% de especificidade, o nosso modelo de GBM previu corretamente o delírio em doentes com menos de 22 anos, com 46 dos 114 verdadeiros positivos (40,4%) em doentes com menos de 65 anos, sugerindo que o nosso modelo está a prever com precisão o delírio, mesmo em populações mais jovens do que os tradicionalmente estudados.

**Conclusões**

Este estudo demonstra a viabilidade da previsão precisa do risco de delírio de incidentes a partir de dados de hospitalização de rotina disponíveis na EHR dentro de 24 horas após a admissão e fornece uma lista de variáveis putativas relacionadas com o delírio que outras instituições podem utilizar para desenvolver os seus próprios modelos. Tal modelo pode permitir uma orientação mais precisa dos recursos de prevenção do delírio para os doentes susceptíveis de serem mais beneficiados.

Pontos-chave:

Pergunta: A aprendizagem mecânica pode ser utilizada para prever delírios de incidentes em pacientes recentemente hospitalizados, utilizando apenas dados disponíveis no registo de saúde electrónico pouco tempo após a admissão?

Descobertas: Neste estudo de coorte, os modelos de classificação foram treinados utilizando 5 algoritmos diferentes de aprendizagem mecânica em 14227 internamentos hospitalares e validados num conjunto de testes prospectivos de 3996 internamentos hospitalares. A máquina de aumento do gradiente modelo com melhor desempenho, com uma área sob a curva característica de funcionamento do receptor de 0,855.

Significado: A aprendizagem da máquina pode prever com precisão o risco de delírio usando dados de registos de saúde electrónicos sobre a admissão e superam as regras de previsão administradas por enfermeiras atualmente utilizadas.