

ESTUDO ORIGINAL

Modelo preditivo para laringoscopia difícil usando aprendizado de máquina: estudo de coorte retrospectivo

Jong Ho Kim^{a,b}, Jun Woo Choi^a, Young Suk Kwon^{a,b,*}, Seong Sik Kang^c

^a Chuncheon Sacred Heart Hospital, Department of Anesthesiology and Pain Medicine, Chuncheon, South Korea

^b Hallym University, Institute of New Frontier Research Team, Chuncheon, South Korea

^c Kangwon National University, College of Medicine, Department of Anesthesiology and Pain Medicine, Chuncheon, South Korea

Recebido em 22 de julho de 2020; aceito em 20 de junho de 2021.

PALAVRAS-CHAVE:

Intubação intratraqueal;
Laringoscópios;
Aprendizado de máquina

RESUMO:

Justificativa: Tanto as previsões quanto os preditores de laringoscopia difícil são controversos. O aprendizado de máquina é um excelente método alternativo para prever laringoscopias difíceis. Este estudo teve como objetivo desenvolver e validar modelos preditivos práticos para laringoscopia difícil por meio de aprendizado de máquina.

Métodos: As variáveis para predição de laringoscopia difícil incluíram idade, grau de Mallampati, índice de massa corporal, distância esternomentoniana e circunferência do pescoço. A laringoscopia difícil foi definida como grau 3 e 4 pela classificação de Cormack-Lehane. Foram incluídos dados pré-anestésicos e anestésicos de 616 pacientes submetidos à anestesia em um único centro. O conjunto de dados foi dividido em um conjunto base de treinamento (n = 492) e um conjunto base de teste (n = 124), com distribuição igual de laringoscopia difícil. Os conjuntos de dados de treinamento foram treinados com seis algoritmos (perceptron multicamada, regressão logística, máquina de vetor de suporte, floresta aleatória, aumento de gradiente extremo e máquina de aumento leve de gradiente) e validação cruzada. O modelo com maior área sob a curva característica de operação do receptor (AUROC) foi escolhido como modelo final, que foi validado com o conjunto de testes.

Resultados: Os resultados da validação cruzada foram melhores usando o algoritmo da máquina de aumento leve do gradiente com grau de Mallampati x idade e distância esternomentoniana como parâmetros do modelo preditivo. O AUROC previsto para a classe de laringoscopia difícil foi de 0,71 (intervalo de confiança de 95%, 0,59-0,83; p = 0,014) e a recordação (sensibilidade) foi de 0,85.

Conclusão: A previsão de laringoscopia difícil é possível com três parâmetros. Danos graves resultantes da falha em prever laringoscopia difícil com alta evocação são pequenos com o modelo relatado. O desempenho do modelo pode ser aprimorado ainda mais por treinamento de dados adicional.

Autor correspondente:

E-mail: gettys@hallym.or.kr (Y.S. Kwon).

<https://doi.org/10.1016/j.bjane.2021.06.016>

© 2021 Sociedade Brasileira de Anestesiologia. Published by Elsevier Editora Ltda. This is an open access article under the CC BY-NC-ND licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Introdução

Via aérea difícil (VAD) é uma situação clínica em que um anesthesiologista tradicionalmente treinado apresenta dificuldade em ventilar a máscara facial da via aérea superior, intubação ou ambos.¹ A incidência de laringoscopia difícil, intubação difícil e falha na intubação foi de 12,3%, 9 % e 0,47%, respectivamente.² Embora as vias aéreas difíceis sejam incomuns em anestesia ou ressuscitação, o manejo inadequado das vias aéreas pode resultar em desfechos fatais. Portanto, prever VAD em pacientes que necessitam de manejo das vias aéreas torna-se crucial. A VAD inclui a laringoscopia (LD) difícil, caso em que a visualização de qualquer porção das pregas vocais não é possível.¹ Existem muitos métodos e diretrizes para avaliar e prever a LD;^{1,3-8} no entanto, ainda há debate sobre os melhores métodos e marcos anatômicos para prever a LD.⁹

O aprendizado de máquina (AM) é uma categoria de algoritmos que permite que aplicativos de software precisos prevejam resultados. O aprendizado de máquina é cada vez mais usado para diagnósticos médicos e previsões de doenças. A premissa básica do aprendizado de máquina envolve escrever um algoritmo que recebe dados de entrada, usa análise estatística para prever uma saída e a atualiza à medida que novos dados se tornam disponíveis. O AM tem o poder de sintetizar e analisar rapidamente combinações multivariáveis complexas. Graças ao AM, a previsão de intubação difícil também está mudando.¹⁰

Existem vários modelos preditivos para LD ou VAD. No entanto, o uso de múltiplos preditores reduz a praticidade,¹⁰⁻¹⁴ e a maioria dos modelos carece de validação suficiente.^{10,11} Este estudo teve como objetivo desenvolver e validar modelos preditivos práticos para EAD por meio de AM usando apenas alguns preditores.

Material e métodos

Este estudo de coorte retrospectivo foi aprovado pelo Conselho de Revisão Institucional/Comitê de Ética do Chuncheon Sacred Heart Hospital, Hallym University (IRB No. 2019-07-012-001). A exigência de consentimento informado por escrito foi dispensada pelo comitê de ética. O estudo foi realizado de forma retrospectiva, e o diagnóstico de LD foi considerado preditor e avaliado comparando-o com o diagnóstico de registro de anestesia. Dados de pacientes que foram submetidos a cirurgia sob anestesia no Hallym University Chuncheon Sacred Heart Hospital entre 18 de janeiro de 2019 e 6 de novembro de 2019 foram coletados dos registros de pré-anestésica e anestesia. Os critérios de exclusão foram: < 18 anos, anestesia regional, grandes anomalias externas faciais ou cervicais, anomalias ou tumores laríngeos, máscara laríngea ou ventilação com máscara, uso de videolaringoscópios ou escopo de fibra óptica para intubação traqueal, intubação realizada por um residente com < 2 anos de experiência em anesthesiologia,

registros insuficientes e estados de intubação endotraqueal antes da anestesia. Dos 3.676 pacientes selecionados, 3.060 foram excluídos e 616 foram finalmente incluídos no estudo.

Parâmetros para predição de laringoscopia difícil

Os parâmetros para predição de LD incluíram idade, índice de massa corporal (IMC), grau de Mallampati, distância esternomentoniana (DEM) e circunferência do pescoço (CP). O grau de Mallampati classifica a facilidade de intubação endotraqueal como Glote classe I – glote totalmente exposta incluindo as comissuras anterior e posterior; Glote classe 2 – glote parcialmente exposta e comissura anterior não visualizada; Glote classe 3 – glote não exposta e apenas cartilagens corniculadas visualizadas; e Glote classe 4 – cartilagens corniculadas não expostas. A DEM foi definida como a distância reta entre a borda superior do manúbrio do esterno e o ponto ósseo do mento após a extensão total do pescoço com a boca fechada.⁸ A circunferência do pescoço foi medida em cm ao nível da cartilagem tireóide.⁷ Todos os valores ausentes foram removidos.

Grau de intubação e laringoscopia

Os procedimentos de rotina para intubação traqueal são padronizados em nosso hospital. Caso a prega vocal não fosse visualizada na primeira tentativa, o anesthesiologista tentaria novamente a manipulação do laringoscópio. Então, dependendo da situação, uma manipulação do laringoscópio foi tentada por outro anesthesiologista ou um videolaringoscópio foi tentado. A laringoscopia difícil foi definida quando todas as tentativas de laringoscopia direta não visualizaram as cordas vocais. No registro de anestesia, se forem feitas várias tentativas, o número de tentativas é registrado e a nota final de Cormack-Lehane é registrada. Lâminas de laringoscópio metálicas descartáveis de uso único Macintosh padrão (INT; Intubrite Llc, Vista, CA, EUA) foram empregadas. No total, sete anesthesiologistas assistentes e três anesthesiologistas residentes classificaram as incidências de laringoscopia direta de acordo com os graus de Cormack-Lehane como Grau 1 = a maior parte da abertura glótica é visível; Grau 2 = apenas a porção posterior da glote ou apenas as cartilagens aritenoides são visíveis; Grau 3 = apenas a epiglote, mas nenhuma porção da glote é visível; Grau 4 = nem a glote nem a epiglote são visíveis. Cormack-Lehane 3 e 4 indicaram LD e foram combinados na classe difícil. Cormack-Lehane 1 e 2 foram combinados na classe não difícil.

Aprendizagem desequilibrada

Os cinco parâmetros de predição mencionados foram incluídos como variáveis. Um conjunto de dados foi criado com as variáveis de 616 pacientes. No entanto, a classe difícil (n = 64) incluiu menos pacientes do que a classe não difícil (n = 552), resultando em dados desbalanceados. Para resolver esse problema, combinamos

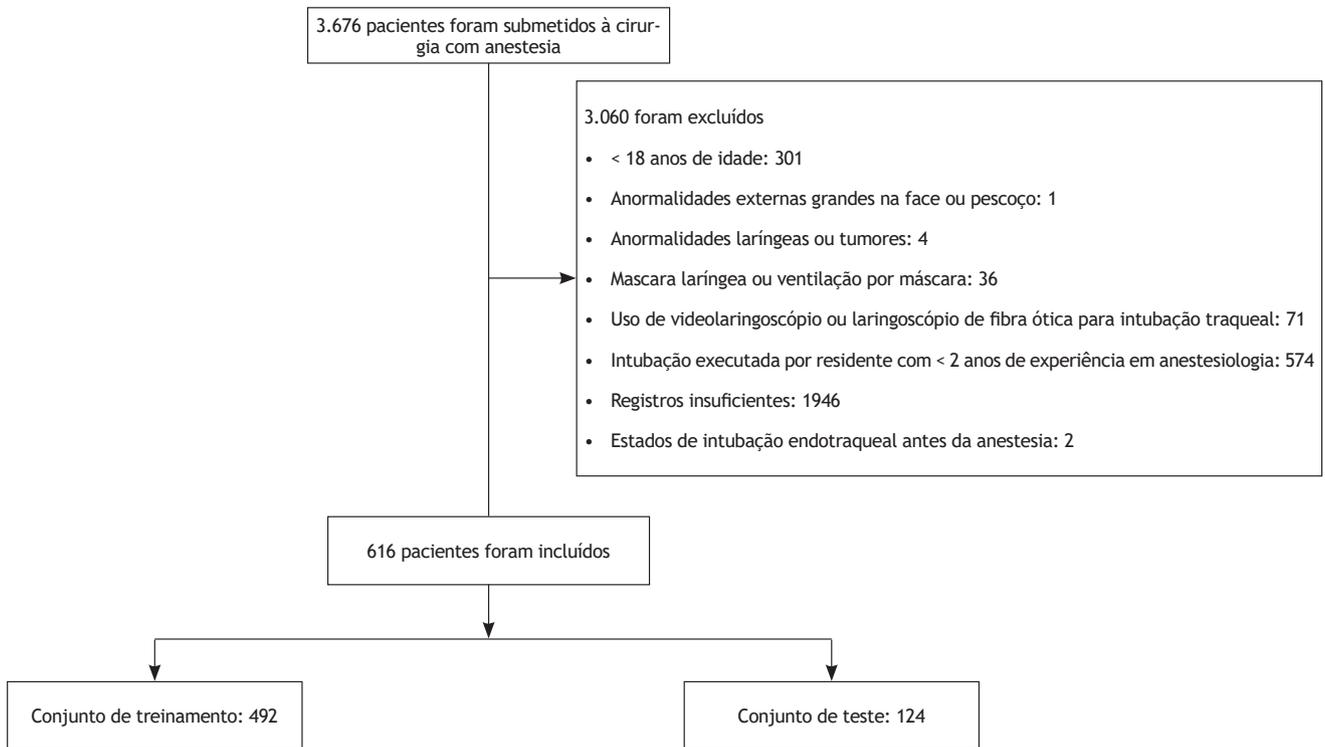


Figura 1 Diagrama de fluxo

super e subamostragem pela técnica de sobreamostragem minoritária sintética (SMOTE).15, 16

Variáveis e seleção de algoritmos

O conjunto de dados foi dividido aleatoriamente em um conjunto de treinamento básico (80%, 492) e um conjunto de teste básico (20%, 124) com distribuição igual de dados de pacientes de classe difícil. Por meio de regressão logística do conjunto de treinamento base, encontramos variáveis com razões de possibilidades significativas ($p < 0,05$) para ocorrência de LD e criamos vários conjuntos de dados de treinamento com diferentes combinações de variáveis. Usamos o conjunto de teste base para criar conjuntos de teste correspondentes às variáveis em cada conjunto de treinamento. Como o objetivo deste estudo é desenvolver um modelo preditivo com um número pequeno de variáveis e bom desempenho, é necessário um modelo treinado com o conjunto básico de treinamento para comparar com um modelo com menos variáveis do que o conjunto básico de treinamento. Além disso, um conjunto de teste correspondente é necessário para avaliar cada conjunto de treinamento, incluindo o conjunto de treinamento básico. Cada conjunto de treinamento foi normalizado pelo escalonamento mín-máx após a aplicação do SMOTE, enquanto cada conjunto de teste foi normalizado pelo escalonamento mín-máx do conjunto de treinamento. Todos os conjuntos de treinamento foram

treinados com seis algoritmos, incluindo perceptron multicamada, regressão logística, máquinas de vetor de suporte, floresta aleatória balanceada, aumento de gradiente extremo e máquinas de aumento leve de gradiente (MALG). Os modelos de predição desenvolvidos aplicaram conjuntos de treinamento aos algoritmos, após os quais foi realizada uma validação cruzada de 10 vezes e calculada a área sob a curva da curva característica de operação do receptor (AUROC). Os resultados foram avaliados pelo AUROC médio, sendo escolhido o melhor modelo como o final.

Avaliação e estatísticas do modelo

O modelo final foi concluído após o ajuste dos hiperparâmetros do modelo selecionado. O desempenho do modelo completo foi avaliado por meio de um conjunto de testes separado e seu desempenho avaliado pelo AUROC. Além disso, foram calculadas a precisão (valor preditivo positivo, razão de pacientes com LD verdadeiro entre os pacientes preditos como LD) e recall (sensibilidade, razão de pacientes preditos como LD entre LD verdadeiro). O subdiagnóstico é LD, mas é previsto e diagnosticado como LND. Se a LD levar à falha no manejo das vias aéreas, pode ter sérias consequências para o paciente. Por outro lado, o sobrediagnóstico está prevendo e diagnosticando o LND como LD e, como o profissional reconhece que o manejo das vias aéreas pode ser difícil e se prepara para isso, o investimento

Tabela 1 Variáveis de predição de laringoscopia difícil em conjuntos de base de dados de treino e teste

	Conjunto de base de dados de treino (n = 492)			Conjunto de base de dados de teste (n = 124)		
	LND(n = 441)	LD(n = 51)	Valor -p	LND(n = 111)	LD(n = 13)	Valor -p
Idade (anos, média ± DP)	52,7 ± 16,4	57,6 ± 14,9	0,044	54,2 ± 16,9	68,5 ± 10,8	0,004
Grau Mallampati (número, %)			0,008			0,057
Grau1	97 (22,0)	7 (13,7)		30 (27,0)	2 (15,4)	
Grau 2	167 (37,9)	13 (25,5)		46 (41,4)	3 (23,1)	
Grau 3	115 (26,1)	19 (37,3)		23 (20,7)	5 (38,5)	
Grau 4	62 (14,1)	12 (23,5)		12 (10,8)	3 (23,1)	
DEM (cm, anos, média ± DP)	17,5 ± 2,3	16,6 ± 1,9	0,006	17,3 ± 2,0	16,4 ± 1,5	0,103
CP (cm, anos, média ± DP)	37,1 ± 3,6	37,3 ± 3,3	0,628	37,3 ± 5,1	37,8 ± 3,1	0,717
IMC	25,5 ± 3,9	25,3 ± 4,0	0,714	25,6 ± 4,0	25,6 ± 4,6	0,723

LND, classe de laringoscopia não difícil; LD, classe de laringoscopia difícil; DEM, distância esternomentoniana; CP, circunferência do pescoço; IMC, índice de massa corporal; DP, desvio padrão.

Tabela 2 Razões de possibilidades das variáveis no conjunto de base de dados de treinamento

	Razões de possibilidades (95% IC)	Valor - p
Idade	1,02 (1,00-1,04)	0,046
Grau Mallampati		0,049
Grau1	reference	
Grau 2	1,08 (0,42-2,80)	0,876
Grau 3	2,29 (0,92-5,68)	0,074
Grau 4	2,68 (1,00-7,18)	0,050
DEM	0,81 (0,70-0,94)	0,005
CP	1,02 (0,94-1,10)	0,627
IMC	0,99 (0,91-1,06)	0,713

DEM, distância esternomentoniana; CP, circunferência do pescoço; IMC, índice de massa corporal.

Tabela 3 Validação cruzada de dez vezes AUROC após a aplicação de cada conjunto de treinamento para cada algoritmo

	Conjunto de treinamento base	Conjunto de treinamento 1	Conjunto de treinamento 2	Conjunto de treinamento 3	Conjunto de treinamento 4
PMC					
(média ± DP)	0,66 ± 0,03	0,60 ± 0,02	0,62 ± 0,04	0,60 ± 0,03	0,62 ± 0,05
RL					
(média ± DP)	0,69 ± 0,1	0,63 ± 0,08	0,75 ± 0,09	0,71 ± 0,15	0,67 ± 0,19
MVS					
(média ± DP)	0,68 ± 0,09	0,63 ± 0,08	0,75 ± 0,1	0,72 ± 0,18	0,68 ± 0,19
FAE					
(média ± DP)	0,92 ± 0,08	0,90 ± 0,09	0,98 ± 0,03	0,98 ± 0,03	0,98 ± 0,04
AGE					
(média ± DP)	0,86 ± 0,1	0,84 ± 0,13	0,99 ± 0,03	0,95 ± 0,05	0,98 ± 0,04
MALG					
(média ± DP)	0,94 ± 0,08	0,92 ± 0,12	*0,99 ± 0,02	0,96 ± 0,06	0,97 ± 0,03

Conjunto de treinamento base: grau de Mallampati, idade, distância esternomentoniana, índice de massa corporal, circunferência do pescoço.

Conjunto de treinamento 1: grau de Mallampati, idade, distância esternomentoniana.

Conjunto de treinamento 2: grau de Mallampati x idade, distância esternomentoniana.

Conjunto de treinamento 3: grau de Mallampati x distância esternomentoniana, idade.

Conjunto de treinamento 4: grau de Mallampati, distância esternomentoniana x idade.

AUROC, área sob a curva característica de operação do receptor; PMC, perceptron multicamada; RL, regressão logística; MVS, máquina de vetores de suporte; FAE, floresta aleatória equilibrada; AGE, aumento de gradiente extremo; MALG, máquina de aumento leve de gradiente.

* O modelo aplicando o conjunto de treinamento 2 à MALG apresentou o melhor desempenho na validação cruzada.

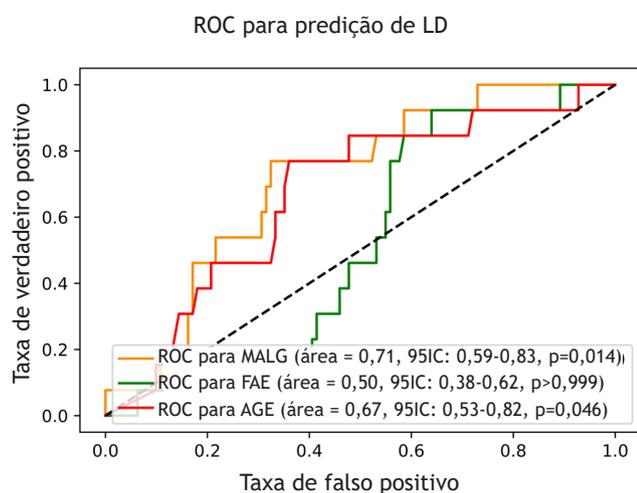


Figura 2 AUROC para algoritmos de modelo de predição de laringoscopia difícil: máquina de aumento leve de gradiente (MALG), aumento de gradiente extremo (AGE), floresta aleatória equilibrada (FAE). Preditores: grau de Mallampati x idade, distância esternomentoniana (DEM). AUROC, área sob a curva característica de operação do receptor; LD, laringoscopia difícil; LND, laringoscopia não difícil; 95CI, intervalo de confiança de 95%.

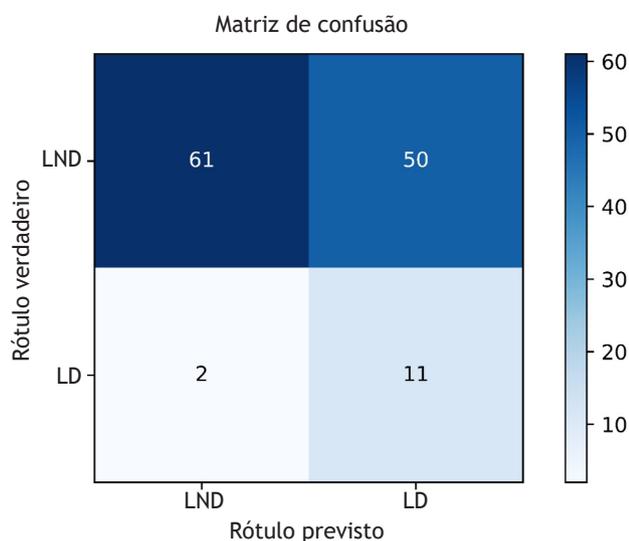


Figura 3 Matriz de confusão do modelo de predição de laringoscopia difícil. LD, laringoscopia difícil; LND, laringoscopia não difícil.

em tempo, mão de obra e equipamentos devido ao superdiagnóstico pode ser desperdiçado.

Todos os dados foram processados e analisados pelo Anaconda (Python versão 3.7; <https://www.anaconda.com>). As variáveis do conjunto de dados foram analisadas pelo SPSS (versão 26.0, IBM). As variáveis contínuas foram comparadas com o teste t ou Mann-Whitney e as variáveis categóricas com o teste qui-quadrado. A razão de chances das variáveis foi determinada por regressão logística (método: enter).

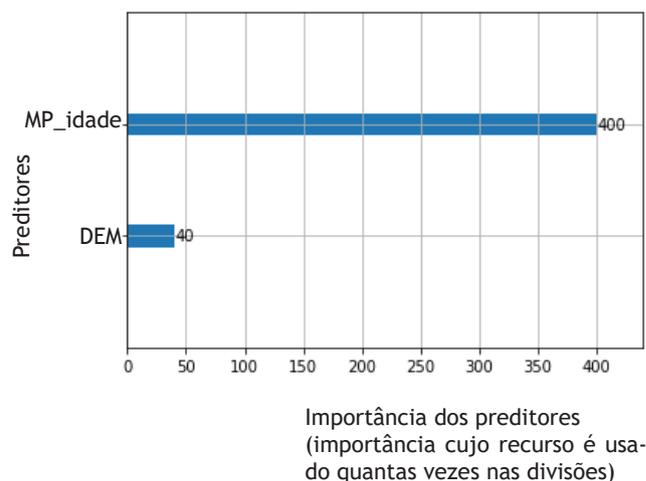


Figura 4 Importância dos preditores no último modelo. MP_idade, grau de Mallampati x idade; DEM, distância esternomentoniana.

Resultados

A Figura 1 mostra a matrícula, o andamento e a alocação do estudo. A Tabela 1 mostra os parâmetros de previsão de LD do paciente no conjunto de dados de treinamento básico e teste básico. A Tabela 2 mostra a razão de chances das variáveis LD no conjunto de treinamento base. Idade, grau de Mallampati e DEM apresentaram significância estatística. Um conjunto de treinamento separado foi criado a partir do conjunto de treinamento básico usando as três variáveis estatisticamente significativas. Além disso, ao combinar as variáveis, foram criadas novas variáveis potenciais e conjuntos de treinamento baseados nas novas variáveis de combinação. As variáveis incluídas no conjunto de dados foram as seguintes:

- Conjunto de treinamento básico e conjunto de teste básico: grau de Mallampati, idade, DEM, IMC, CP
- Conjunto de treinamento 1 e conjunto de teste 1: grau de Mallampati, idade, DEM
- Conjunto de treinamento 2 e conjunto de teste 2: grau de Mallampati x idade, DEM
- Conjunto de treinamento 3 e conjunto de teste 3: grau de Mallampati x DEM, idade
- Conjunto de treinamento 4 e conjunto de teste 4: grau de Mallampati, DEM x idade

Os resultados da validação cruzada após a aplicação de cada conjunto de trens para cada algoritmo são mostrados na Tabela 3. O desempenho do modelo foi melhor quando Mallampati grau x idade e DEM foram aplicados à MALG na validação cruzada. Portanto, selecionamos o modelo aplicando as variáveis de Mallampati grau x idade e DEM para MALG como modelo de predição para LD. O modelo final foi concluído após

o ajuste dos hiperparâmetros. O conjunto de teste foi aplicado ao modelo final para avaliar o desempenho do modelo final, que apresentou um AUROC de 0,71 (intervalo de confiança de 95%, 0,59–0,83; $p = 0,014$) (Figura 2). A matriz de confusão mostra os resultados da contagem do número de correspondências entre a classe original do conjunto de teste alvo e a classe do conjunto de teste prevista pelo modelo (Figura 3). A precisão para a predição de LD foi de 0,18 (proporção de pacientes com DL verdadeiros entre pacientes preditos como LD, 11/60) e o recall de 0,85 (proporção de pacientes preditos como LD entre pacientes com LD verdadeira, 11/13). A importância dos preditores usados no modelo final é mostrada na Figura 4.

Discussão

Numerosos estudos relataram várias combinações de parâmetros e resultados na tentativa de prever a intubação da VAD. Em uma metanálise prevendo intubação difícil, a combinação do grau de Mallampati e da distância tireoentoniana predisse a intubação mais difícil,¹⁷ entretanto, a falta de heterogeneidade e o pequeno número de estudos limitaram as conclusões. L’Hermite et al. previram intubações difíceis usando os escores passivos de cinco parâmetros das vias aéreas, mas sugeriram que tais predições de intubação difícil provavelmente não seriam úteis.¹⁸ Alguns estudos usaram AM para prever LD ou intubação difícil. No entanto, os resultados desses estudos não foram validados e avaliados usando um conjunto de testes. Moustafa et al. usaram ML com valor preditivo positivo de 76%, valor preditivo negativo de 76% e AUROC de 0,79 para LD.¹¹ No entanto, mesmo com nove parâmetros preditivos, seus resultados de validação cruzada foram piores do que o estudo atual (AUROC, 0,79 vs. 0,99). Em seu estudo, Langeson et al. concluíram que, embora a intubação difícil usando cálculo assistido por computador fosse preditiva, cada parâmetro usado para intubação difícil tinha um valor preditivo ruim.¹⁰ Seu alvo preditivo final era intubação difícil, que difere ligeiramente de nosso alvo preditivo – predição de LD com menos variáveis. Eles relataram variáveis significativas no IMC, idade e grau de Mallampati, mas em nosso estudo, o resultado da validação cruzada de modelos sem IMC foi melhor.

Apesar das diferenças nos alvos de predição, alguns modelos de predição de intubação traqueal difícil que não empregam ML têm um desempenho melhor do que o nosso. Os resultados validados para esses modelos variam de 0,79 a 0,87 no AUROC.¹¹ No entanto, os modelos anteriores usavam de 4 a 9 preditores.¹⁰⁻¹⁴ O modelo de Arne empregou sete variáveis e apresentou o maior AUROC entre os modelos validados.¹³ Embora muitos preditores possam ser usado para previsões precisas, aumentar seu número não é prático. Neste estudo, o objetivo foi criar modelos preditivos que tenham bom desempenho com poucos preditores. A previsão de vias aéreas com poucos parâmetros é um bom alvo

para o AM devido ao seu poder de sintetizar e analisar rapidamente combinações de variáveis. Por outro lado, modelos com alta precisão e recall são ideais para predição de LD. A alta precisão durante a previsão de LD pode economizar tempo e dinheiro na preparação de VAD. O recall é ainda mais crítico considerando o sério problema de LD imprevisível. No entanto, quando o conjunto de teste foi aplicado ao nosso modelo, ele não apresentou o mesmo resultado que os resultados da validação cruzada, e a precisão foi baixa. Isso pode ser atribuído ao overfitting dos dados de treinamento. Em geral, a solução para o overfitting é coletar mais dados de treinamento.¹⁹ De fato, a pequena proporção de pacientes com LD não forneceu um número suficiente de conjuntos de dados. Nesses casos de dados desbalanceados, a quantidade de dados se torna ainda mais crítica.

Em nosso modelo, os resultados da validação cruzada melhoraram apesar da redução do número de preditores. Além disso, os resultados da validação cruzada do modelo usando o recém-combinado grau de Mallampati x idade foram melhores do que aqueles com grau de Mallampati e idade como preditores separados. Nas previsões de AM, o desempenho do modelo pode ser melhorado removendo informações irrelevantes para previsões ou combinando recursos para criar recursos potencialmente novos.¹⁹

Mostramos um modelo que prevê LD com o menor número de preditores possível, podendo identificar preditores importantes para classificar LDs em modelos de aprendizado de máquina. No entanto, é difícil interpretar como esses preditores se comportam em modelos de aprendizado de máquina. O aprendizado de máquina é frequentemente chamado de caixa preta. Os dados entram e as decisões saem, mas o processo entre entrada e saída é opaco. número.

A intubação difícil muitas vezes pode levar a consequências inesperadas e complexas, prejudicando o paciente. Em 2015–2018, houve diferença nos custos de hospitalização nos Estados Unidos entre pacientes com intubação de via aérea difícil e aqueles que não o fizeram, e o valor médio foi cerca de US\$ 20.000 maior em pacientes com intubação de via aérea difícil.²¹ Laringoscopia difícil é diferente de difícil intubação traqueal, mas pode causar intubação traqueal difícil. Em alguns estudos, metade das laringoscopias difíceis relatou uma escala de dificuldade de intubação alta ≥ 5 pontos.²² A predição adequada da laringoscopia difícil pode ajudar a antecipar vias aéreas difíceis e ser uma maneira eficaz de reduzir os custos médicos.

Nosso estudo tem algumas limitações. Primeiro, semelhante a estudos anteriores de aprendizado de máquina, não se aplica a pacientes pediátricos, pois se aplica a populações adultas. Segundo, porque tanto o treinamento quanto os testes foram conduzidos em indivíduos asiáticos, os resultados podem diferir em outras raças. Por exemplo, populações asiáticas têm dimensões estatisticamente diferentes das populações caucasianas para arco do queixo, comprimento da face

e protrusão do nariz.²³ Para generalizar para pacientes com anestesia global, estudos com variáveis adicionais podem ser necessários. Recentemente, os videolarinoscópios têm sido usados em vias aéreas difíceis, de modo que o manejo das vias aéreas difíceis é melhor do que no passado. No entanto, em situações em que os videolarinoscópios não estão disponíveis, a difícil previsão do laringoscópio pode ser útil. Terceiro, o uso de videolarinoscópio ou endoscópio de fibra óptica para intubação traqueal foi excluído em nosso estudo. Reduz a incidência de resultados de laringoscopia difíceis, o que pode limitar a variabilidade dos preditores.

Em conclusão, o novo modelo proposto pode prever LD com apenas três preditores (grau de Mallampati, idade e distância esternomental). Esse pequeno número de preditores torna a previsão simples e com alto recall. Assim, a probabilidade de problemas sérios causados por falha na previsão de LD é baixa. No entanto, para melhorar o desempenho geral e a generalização do modelo, são necessários mais dados de pacientes com LDs e de diferentes raças.

Conflitos de interesse

Os autores declaram não haver conflitos de interesse.

Financiamento

Esta pesquisa foi financiada pelo First Research in Lifetime Program da National Research Foundation (NRF) financiado pelo governo coreano (MSIT) (NRF-2018R-1C1B5085866), Coreia do Sul.

References

1. Apfelbaum JL, Hagberg CA, Caplan RA, Blitt CD, Connis RT, Nickinovich DG, et al. Practice guidelines for management of the difficult airway: an updated report by the American Society of Anesthesiologists task force on management of the difficult airway. *Anesthesiology*. 2013;118(2):251-70. <https://doi.org/10.1097/ALN.0b013e31827773b2>
2. Workeneh SA, Gebregzi AH, Denu ZA. Magnitude and Predisposing Factors of Difficult Airway during Induction of General Anaesthesia. *Anesthesiol Res Pract*. 2017;2017:5836397-. <https://doi.org/10.1155/2017/5836397>
3. Etezadi F, Ahangari A, Shokri H, Najafi A, Khajavi MR, Daghigh M, et al. Thyromental height: a new clinical test for prediction of difficult laryngoscopy. *Anesth Analg*. 2013;117(6):1347-51. <https://doi.org/10.1213/ANE.0b013e3182a8c734>
4. Frerk C. Predicting difficult intubation. *Anaesthesia*. 1991;46(12):1005-8. <https://doi.org/10.1111/j.1365-2044.1991.tb09909.x>
5. Khan ZH, Kashfi A, Ebrahimkhani E. A comparison of the upper lip bite test (a simple new technique) with modified Mallampati classification in predicting difficulty in endotracheal intubation: a prospective blinded study. *Anesth Analg*. 2003;96(2):595-9. <https://doi.org/10.1097/0000539-200302000-00053>
6. Mallampati SR, Gatt SP, Gugino LD, Desai SP, Waraksa B, Freiburger D, et al. A clinical sign to predict difficult tracheal intubation; a prospective study. *Can Anaesth Soc J*. 1985;32(4):429-34. <https://doi.org/10.1007/BF03011357>.
7. Riad W, Vaez MN, Raveendran R, Tam AD, Quereshy FA, Chung F, et al. Neck circumference as a predictor of difficult intubation and difficult mask ventilation in morbidly obese patients: A prospective observational study. *Eur J Anaesthesiol*. 2016;33(4):244-9. <https://doi.org/10.1097/EJA.0000000000000324>.
8. Savva D. Prediction of difficult tracheal intubation. *Br J Anaesth*. 1994;73(2):149-53. <https://doi.org/10.1093/bja/73.2.149>
9. Türkan S, Ates Y, Cuhruk H, Tekdemir I. Should we re-evaluate the variables for predicting the difficult airway in anesthesiology? *Anesth Analg*. 2002;94(5):1340-4. <https://doi.org/10.1097/0000539-200205000-00055>
10. Langeron O, Cuvillon P, Ibanez-Esteve C, Lenfant F, Riou B, Le Manach Y. Prediction of Difficult Tracheal Intubation Time for a Paradigm Change. *Anesthesiology*. 2012;117(6):1223-33. <https://doi.org/10.1097/ALN.0b013e31827537cb>
11. Moustafa MA, El-Metainy S, Mahar K, Mahmoud Abdel-magied E. Defining difficult laryngoscopy findings by using multiple parameters: A machine learning approach. *Egyptian Journal of Anaesthesia*. 2017;33(2):153-8. <https://doi.org/10.1016/j.egja.2017.02.002>
12. Wilson M, Spiegelhalter D, Robertson J, Lesser P. Predicting difficult intubation. *Br J Anaesth*. 1988;61(2):211-6. <https://doi.org/10.1093/bja/61.2.211>
13. Arne J, Descoins P, Fuscuardi J, Ingrand P, Ferrier B, Boudigues D, et al. Preoperative assessment for difficult intubation in general and ENT surgery: predictive value of a clinical multivariate risk index. *Br J Anaesth*. 1998;80(2):140-6. <https://doi.org/10.1093/bja/80.2.140>
14. Naguib M, Malabarey T, AlSatli RA, Al Damegh S, Samarkandi AH. Predictive models for difficult laryngoscopy and intubation. A clinical, radiologic and three-dimensional computer imaging study. *Can J Anaesth* 1999;46(8):748. <https://doi.org/10.1007/BF03013910>
15. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*. 2002;16:321-57. <https://doi.org/10.1613/jair.953>
16. Wilson DL. Asymptotic properties of nearest neighbor rules using edited data. *IEEE Transactions on Systems, Man, and Cybernetics*. 1972(3):408-21. <https://doi.org/10.1109/TSMC.1972.4309137>
17. Shiga T, Wajima Zi, Inoue T, Sakamoto A. Predicting Difficult Intubation in Apparently Normal Patients: A Meta-analysis of Bedside Screening Test Performance. *Anesthesiology*. 2005;103(2):429-37. <https://doi.org/10.1097/0000542-200508000-00027>
18. Yentis SM. Predicting difficult intubation--worthwhile exercise or pointless ritual? *Anaesthesia*. 2002;57(2):105-9. <https://doi.org/10.1046/j.0003-2409.2001.02515.x>
19. Géron A. Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems: 1st ed. Sebastopol (CA): O'Reilly Media, Inc. 2017.
20. Opening the black box of machine learning, [https://www.thelancet.com/journals/lanres/article/PIIS2213-2600\(18\)30425-9/fulltext](https://www.thelancet.com/journals/lanres/article/PIIS2213-2600(18)30425-9/fulltext) [accessed 11 January 2011].

21. Phillips K, Moucharite M, Wong T, May R. Cost burden associated with difficult intubation in the United States. *Trends in Anaesthesia and Critical Care*. 2020;30:e131.
22. Adnet F, Borron SW, Racine SX, Clemessy J-L, Fournier J-L, Plaisance P, et al. The intubation difficulty scale (IDS) proposal and evaluation of a new score characterizing the complexity of endotracheal intubation. *Anesthesiology*. 1997;87(6):1290-7. <https://doi.org/10.1097/00000542-199712000-00005>.
23. Zhuang Z, Landsittel D, Benson S, Roberge R, Shaffer R. Facial anthropometric differences among gender, ethnicity, and age groups. *Ann Occup Hyg*. 2010;54(4):391-402. <https://doi.org/10.1093/annhyg/meq007>.