

Evaluation of a general emergency dataset for the prediction of videolaryngoscopy use in emergency anaesthesia in prehospital trauma care – a proof of concept with consequences?

A. Luckscheiter¹ · W. Zink¹ · J. Eisenberger² · T. Viergutz³ · M. Thiel⁴

► **Zitierweise:** Luckscheiter A, Zink W, Eisenberger J, Viergutz T, Thiel M: Evaluation des Minimalen Notfalldatensatzes zur Vorhersage der Verwendung der Videolaryngoskopie bei der Prähospitalnarkose des Polytraumapatienten – Ein Proof of Concept mit Folgen? Anästh Intensivmed 2024;65:357–367. DOI: 10.19224/ai2024.357

- 1 Klinik für Anästhesiologie, Operative Intensivmedizin und Notfallmedizin, Klinikum Ludwigshafen (Direktor Prof. Dr. W. Zink)
- 2 Stelle zur trägerübergreifenden Qualitätssicherung im Rettungsdienst Baden-Württemberg (SQR-BW), Stuttgart (Leitung: Dr. J. Koster)
- 3 Klinik für Anästhesiologie, Intensivmedizin und Schmerztherapie, Berufsgenossenschaftliche Unfallklinik Tübingen (Direktor: Prof. Dr. T. Viergutz)
- 4 Klinik für Anästhesiologie und Intensivmedizin, Universitätsklinikum Mannheim (Direktorin: Prof. Dr. G. Beck)

Danksagung

Die Autoren möchten Dr. Thorsten Lohs für die Unterstützung bei der Realisierung dieses Forschungsvorhabens danken.

Anmerkung

Im Manuskript wird aus Gründen der Lesbarkeit auf einen Genderstern verzichtet und das grammatikalisch korrekte Genus verwendet. Sofern nicht explizit vermerkt, sind immer alle Geschlechter gleichermaßen damit gemeint.

Interessenkonflikt

Die Autorinnen und Autoren geben an, dass keine Interessenkonflikte bestehen.

Schlüsselwörter

Atemwegsmanagement – Klinische Entscheidungsregeln – Entscheidungsbäume – Maschinelles Lernen – Videolaryngoskopie

Keywords

Airway Management – Clinical Decision Rules – Decision Trees – Machine Learning – Laryngoscopy, Video-assisted

Zusammenfassung

Hintergrund

Die Anwendung der Videolaryngoskopie (VL) ist gemäß aktuellen Leitlinien ein Hauptqualitätsmerkmal der präklinischen Atemwegssicherung bei Traumpatienten. Der zur notärztlichen Dokumentation verwendete Minimale Notfalldatensatz (MIND) soll hinsichtlich der Vorhersehbarkeit der Verwendung der VL validiert werden. Zudem soll die Modellierung mit Blick auf Stabilität bei steigender Nutzerakzeptanz der VL evaluiert werden.

Methodik

Aus einem südwestdeutschen Register wurden Echtweltdaten erwachsener Traumpatienten der Jahre 2018–2020 nach alleiniger notärztlicher Atemwegssicherung, Ausschluss einer Reanimationspflichtigkeit sowie vollständig dokumentiertem Atemwegsmanagement selektioniert. Mittels einer k-Means-Clusteranalyse sowie eines Random-Forest-Modells wurden die MIND-Attribute initiale Vitalwerte, Auskultationsbefunde, notärztliches Erfahrungsniveau in der Atemwegssicherung, Verletzungsmuster, Kapnographie, nichtinvasive Beatmung, supraglottische Atemwegshilfe, Muskelrelaxanzienangabe, Halswirbelsäulenimmobilisation, Thoraxdrainagenanlage und Angaben zum schwierigen Atemweg evaluiert. Um eine steigende Verwendung zu simulieren, wurde vom Originaldatensatz (1:2,6) ausgehend das Verhältnis „VL“ zu „keine VL“ 1:1 bzw. 2:1 randomisiert.

Evaluation des Minimalen Notfalldatensatzes zur Vorhersage der Verwendung der Videolaryngoskopie bei der Prähospitalnarkose des Polytraumapatienten – Ein Proof of Concept mit Folgen?

Ergebnisse

Aus 1.451 Prähospitalnarkosen konnten 992 ausgewertet werden (277-malige Verwendung der VL (28 %)). Mit simulierter steigender Verwendung der VL steigerte sich die Sensitivität von 0,66 (Original) auf 0,77 (2:1) bzw. der positiv prädiktive Wert von 0,5 (Original) auf 0,82 (2:1). Die Spezifität fiel von 0,82 (Original) auf 0,58 (2:1) bzw. der negativ prädiktive Wert von 0,9 (Original) auf 0,51 (2:1) ab. Die Fläche unter der Grenzwertoptimierungskurve variierte von 0,78 (Original) bis 0,77 (2:1). Der Cluster „VL“ enthielt die Attribute Spastik, Dyspnoe oder Zyanose, nichtinvasive Beatmung, Alter, Glasgow Coma Scale und Vitalwerte.

Schlussfolgerung

Auf Basis des MIND scheint nicht nur eine Modellierung der Verwendung der VL bei Traumpatienten möglich, sondern auch eine Unterteilung nach Anwendungscharakteristika. Eine Erweiterung um zusätzliche Attribute wie Gerätetypen, Atemwegevaluation oder Vorgehen bei der Atemwegssicherung könnte die Robustheit zukünftiger Modelle erhöhen.

Summary

Background

In current German guidelines, the use of video laryngoscopy is one of the main quality features in the preclinical airway management of patients with severe injuries. The aim of the study is to evaluate the nationwide used emergency

dataset MIND for the prediction of video laryngoscopy use in prehospital trauma patients requiring emergency anaesthesia with the help of machine learning and considering various application frequencies.

Methods

From a southwest German register we selected real-life datasets pertaining to severely injured adults who had received airway management from emergency physicians in the years between 2018 and 2020. With machine learning, the use of video laryngoscopy (VL) in preclinical adult trauma patients was modeled with a random forest algorithm and also clustered with k-means algorithm. Based on the original dataset (ratio 1:2,6), datasets with a 1:1 and a 2:1 ratio VL/no VL were created at random. Attributes chosen were initial vital signs, experience in airway management as well as injury pattern, auscultation findings, capnography, non-invasive ventilation, supraglottic airway devices, muscle relaxant, immobilisation of the cervical spine, declaration of a difficult airway situation and thoracic drainage insertion.

Results

Out of a total of 1,451 emergency anaesthesias, VL was applied in 277 of the 992 cases evaluated (28 %). With simulated growing usage frequency, sensitivity increased from 0.66 (original) up to 0.77 (2:1) and a positive predictive value rose from 0.5 (original) to 0.82 (2:1). The specificity and the negative predictive value decreased from 0.82 (original) to 0.58 (2:1) and from 0.9 (original) to 0.51 (2:1), respectively. The area under the receiver operating characteristic varied from 0.78 (original) to 0.77 (2:1). The cluster "VL" contained the attributes bronchial spasm, dyspnoea or cyanosis, non-invasive ventilation, age, Glasgow Coma Scale (GCS) and vital parameters.

Conclusion

The attributes of MIND could be used to develop a preliminary prediction model for video laryngoscopy usage in the field of prehospital trauma care. Also, clustering revealed a characteristic type of usage. In order to increase sta-

bility, further attributes like device type, airway evaluation or airway management approach could be included in future models.

Einleitung

Schwerverletzte Notfallpatienten weisen eine höhere Inzidenz eines schwierigen Atemwegs aufgrund von Gesichts- und Halswirbelsäulenverletzungen auf und sind auch hinsichtlich Aspiration und Traumata von Atemwegen und Lunge von Hypoxie bedroht [1–3]. Die Rate an Komplikationen steigt zudem mit zunehmender Anzahl an Atemwegssicherungsversuchen. Durch ihr Konstruktionsprinzip gelingt bei primärer Anwendung der Videolaryngoskopie (VL) häufiger und schneller eine Visualisierung der Glottisebene, entweder direkt oder indirekt [4–6]. Daher empfehlen die Leitlinien zur prähospitalen Atemwegssicherung sowie zur Polytraumaversorgung ihre primäre Verwendung, auch um mögliche Atemwegstraumata bei wiederholten konventionellen Intubationsversuchen zu vermeiden [2,3]. Aktuelle Registerdaten und Untersuchungen zeigen aber, dass die VL oftmals von Notärzten eher als Rückfallebene zur klassischen Intubation verwendet wird [7–9]. Faktoren, die den Entscheidungsprozess umfassend darstellen, konnten bisher nicht ausreichend identifiziert werden [7,10].

Maschinelles Lernen hat in der Medizin in den letzten Jahren vermehrt zur Analyse von Entscheidungsprozessen bzw. als Unterstützungstool in Diagnostik und Therapie Eingang gefunden. Durch die Einführung standardisierter elektronischer Dokumentationssysteme in der präklinischen Notfallmedizin haben sich nun auch hier Möglichkeiten für den Einsatz dieser Methode ergeben. In Grundlagenstudien ist es daher nötig, diese Datenbanken zu bewerten. In Deutschland existieren drei große, überregionale prähospitalen Datenbanken: das Reanimationsregister, das Traumaregister sowie die Stelle zur trägerübergreifenden Qualitätssicherung im Rettungsdienst Baden-Württemberg (SQR-BW). Das Traumaregister erfasst

bei Prähospitalnarkosen die Videolaryngoskopie nicht, wohingegen das Reanimationsregister sie im gültigen Erfassungsbogen als Atemwegshilfsmittel ausweist [11,12]. Einzig die SQR-BW erfasst für alle Notfallbilder in ihrer Datenbank mittels des Minimalen Notfalldatensatzes (MIND) ihre Anwendung [8,13]. Wie bereits in einer Untersuchung zur Vorhersage der Atemwegssicherung beim Polytrauma gezeigt, ist der MIND grundsätzlich für die Forschung mit maschinellem Lernen geeignet [14]. Zudem zeigten Untersuchungen mit diesen „Realweltdaten“ bereits verschiedene Strategien der Atemwegssicherung abhängig von Fachrichtung und Weiterbildung der Notärzte, obgleich die Register- und Datenstruktur Vorteile wie auch Limitationen ausweist [9]. Als Vorteile wären die standardisierte Erfassung, direkte Verfügbarkeit, Überregionalität sowie die Registergröße zu nennen, als Nachteile die starre Datenstruktur, die keine Erweiterung bzw. zeitliche Auflösung ermöglicht und Verlaufsparemeter und Freitexte nicht erfasst. Echtweltdaten unterliegen zudem weniger einem Selektionsbias als in einem Studienregister [15].

Die vorliegende Studie möchte daher im Sinne eines ersten Proof of Concept der Frage nachgehen, inwieweit sich der MIND trotz seiner allgemeinen Limitationen für die Vorhersagen der Verwendung der Videolaryngoskopie mittels Algorithmen des maschinellen Lernens bei Polytraumapatienten eignet. Welchen Einfluss hat eine simulierte Erhöhung der Verwendung der Videolaryngoskopie auf die Aussagekraft solch eines Modells? Welche Schlussfolgerungen lassen sich für die Leitlinie zum prähospitalen Atemwegsmanagement, zur rettungsdienstlichen Organisationsstruktur sowie zur Datensatzerweiterung ableiten, sofern die Modelle eine ausreichende Diskriminationsfähigkeit aufweisen?

Methodik

Die vorliegende Studie wird nach dem Bericht "Transparent Reporting of a Multivariable Prediction Model for indivi-

dual Prognosis or Diagnosis" (TRIPOD) wiedergegeben [16]. Die Untersuchung ist eine Sekundärauswertung und wurde ursprünglich von der Ethik-Kommission der Landesärztekammer Rheinland-Pfalz genehmigt (Applikationsnummer 2021-15767-retrospektiv).

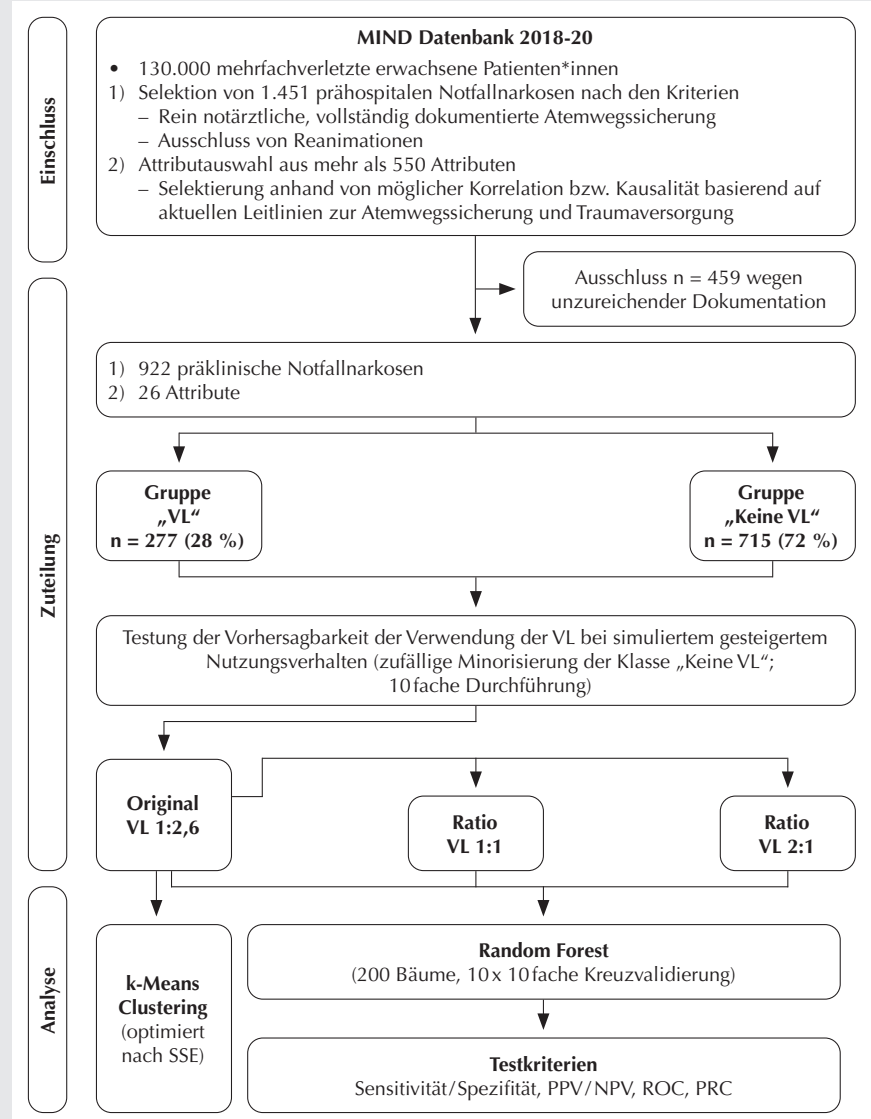
Hintergrund und Studiendesign

Aus dem Register der SQR-BW wurden multipel verletzte erwachsene Patienten der Jahre 2018–2020 selektiert, die prähospital eine Atemwegssicherung erhielten. Hierbei wurden nur Fälle selektiert, bei denen notärztlicherseits eine Atemwegssicherung dokumentiert wurde (und nicht durch das Rettungsdienstpersonal). Reanimationspflichtige Patienten wurden bei Fokussierung auf Notfallnarkosen ausgeschlossen. Ebenso wurden Patienten ohne vollständig dokumentiertes Atemwegsmanagement im Sinne der Fragestellung nicht inkludiert. Als Prämisse wurden dokumentierte Werte als korrekt bzw. tatsächlich durchgeführt gewertet. Eine Interpolation fehlender Werte war bei binären bzw. fehlenden (Verlaufs-)Parametern nicht sinnvoll. Es erfolgte anschließend eine Einteilung in die Klassen „Videolaryngoskopie (VL)“ und „keine Videolaryngoskopie („Keine VL“, Abb. 1). Eine primäre oder sekundäre Verwendung, wie in vorherigen Studien abgeleitet, wurde hier nicht berücksichtigt.

Attributauswahl und Datenvorverarbeitung

Als Grundlage für die Attributauswahl im MIND (hier Version 3.1) dienen die Leitlinie zum präklinischen Atemwegsmanagement sowie Vorgaben der prähospitalen Traumaversorgung [2,17]. Aufgrund vorangegangener Untersuchungen zum Atemwegsmanagement in Abhängigkeit von Erfahrung und Weiterbildung der Notärzte wurde in vorliegender Studie ebenfalls das Dreyfus-Modell der Grundstruktur der Expertise (auch „Modell der fünf Kompetenzstufen“) verwendet, welches an den MIND adaptiert werden kann [9,18]. So galten als Notärzte mit hoher Erfahrung sämtliche anästhesiologischen Fachärzte mit und ohne Zusatzbezeichnung Intensivmedi-

Abbildung 1



Flussdiagramm zu Studienaufbau, Patientenauswahl, Erstellung des Datensets sowie zu Training und Testung.

VL: Videolaryngoskopie; **MIND:** Minimaler Notfalldatensatz; **PPV:** positiv prädiktiver Wert; **NPV:** negativ prädiktiver Wert; **ROC:** Grenzwertoptimierungskurve (receiver operator characteristics); **PRC:** Sensitivitäts-PPV-Kurve (precision recall curve); **SSE:** Summe der quadratischen Fehler.

zin. Alle nicht anästhesiologischen Ärzte mit der Zusatzbezeichnung Intensivmedizin sowie alle anästhesiologischen Ärzte in Weiterbildung wurden der Gruppe mit mittlerem Erfahrungsniveau zugeordnet und alle weiteren Notärzte in der Gruppe mit niedrigem Erfahrungsniveau. Neben dem Verletzungsmuster der verschiedenen Körperregionen (subjektive notärztliche Einteilung in „keine“,

„leicht“, „mittel“ und „schwer“) wurden die initialen Vitalwerte sowie die initiale Glasgow Coma Scale (GCS), die Verwendung von Kapnographie, Muskelrelaxanzien, supraglottischer Atemwegshilfen sowie der nichtinvasiven Beatmung berücksichtigt. Als weitere Interventionen im Zusammenhang mit dem Atemweg bzw. der Atmung wurden die Anlage einer Thoraxdrainage bzw. eine Halswir-

belsäulenimmobilisation als Surrogate für traumabedingte Ateminsuffizienz bzw. erschwerten Atemwegszugang mit aufgenommen. Weitere Attribute waren die Auskultationsbefunde, der Pre-Emergency-Status (PES, eine präklinisch adaptierte ASA-Klassifikation) sowie das Patientenalter. Bezüglich der Angabe eines schwierigen Atemwegs wurden die Attribute schwierige Atemwegsanatomie, mehr als zwei Intubationsversuche und schwierige Maskenbeatmung berücksichtigt.

Modellierung und Algorithmus

Um eine steigende Anwendungshäufigkeit der Videolaryngoskopie zu simulieren, wurden hierfür aus dem Originaldatensatz (Verhältnis „VL“ zu „keine VL“ 1:2,6) zweimalig neue Datensätze im Zufallsverfahren erstellt. Hierzu wurde zehnmal ein Datensatz im Verhältnis „VL/keine VL“ 1:1 bzw. 2:1 randomisiert. Dieses Vorgehen diente auch dazu, die Limitation der unklaren Vorhaltung der VL auf den arztbesetzten Rettungsmitteln zu minimieren. Anschließend erfolgte die Analyse mittels maschinellen Lernens. Als Algorithmus wurde der Random-Forest-Algorithmus gewählt [19]. Random Forest erstellt eine bestimmte Anzahl von Zufallsbäumen (hier: $n = 200$) anhand einer zufälligen Stichprobe (100 % des Trainingsatzes) mit einer zufälligen Anzahl an Attributen, wobei jeder Baum eine Vorhersage betreffend die Fragestellung erstellt. Anschließend wird ein Schwellenwert bestimmt, wie viele Bäume für oder gegen eine Entscheidung votieren müssen, und dementsprechend ein finales Ergebnis ausgegeben. So ergibt sich gegenüber einem einzelnen Baum der Vorteil, dass bei inhomogenen oder sich verändernden Datenstrukturen im Allgemeinen eine erhöhte Robustheit erreicht und so der prognostischen Aussagekraft gerade bei binären Fragestellungen mehr Stabilität verliehen werden kann. Zudem kann der Berechnungsprozess durch die Attributsgewichtung (Methode: mittlere Abnahme der Unreinheit) nachvollzogen werden. Random Forest kann auch mit fehlenden Attributen umgehen [19].

Training, Testung und Leistungsfähigkeit

Die Analyse und Modellierung erfolgten mittels der Software WEKA 3.8.4 (Waikato Environment for Knowledge Analysis, University of Waikato, Neuseeland). Als Gütekriterien wurden Sensitivität und Spezifität, positiver und negativer prädiktiver Wert (PPV, NPV), die Fläche unter der Grenzwertoptimierungskurve (receiver operating characteristic, AUC-ROC) sowie unter der Precision-Recall-Kurve (AUC-PRC, Sensitivität-PPV-Kurve) untersucht [19]. Bei der AUC-ROC wurde gemäß der Einteilung nach Nham eine Fläche unter Kurve von $\geq 0,9$ als exzellent, $0,8-0,9$ als gut, von $0,7-0,8$ als akzeptabel, $0,6-0,7$ als unzureichend und $\leq 0,6$ als schlecht gewertet [20]. Ebenso wurden die Durchschnittswerte von Kappa als Maß für das Verhältnis zwischen gemessener zu zufälliger Übereinstimmung (Reliabilität) und der Matthews-Korrelationskoeffizient (MCC) als Maß für die Güte der binären Klassifikation ermittelt [21]. Zur Validierung wurde ein Kreuzvalidierungsverfahren eingesetzt. Hierbei wird der Datensatz zufällig in zehn gleiche Teile unterteilt. Neun werden für das Training und einer für die Testung verwendet und das Verfahren so oft wiederholt, bis jeder Abschnitt einmal als Testdatensatz genutzt wurde. Dann wird das Modell ausgegeben. Beim Originaldatensatz wurden diese Kreuzvalidierung zehnfach in unterschiedlicher Randomisierung durchgeführt und die Werte anschließend gemittelt. Für die anderen beiden Gruppen wurde für jedes der zehn randomisierten Datensets die Kreuzvalidierung einmalig durchgeführt und dann die Werte gemittelt. Hierdurch kann eine breitere Varianz in der Modellierung abgedeckt, jedoch kein finales Modell ausgegeben werden. Der statistische Vergleich der Gütekriterien bzw. der Attribute erfolgte mittels T- oder Chi-Quadrat-Test ($p < 0,05$ als signifikant) in Microsoft Excel 2016 (Microsoft Corporation, Redmond, Seattle, USA).

Clustering

Unter Clustering versteht man die Ähnlichkeitsanalyse mittels mathematischer

Verfahren. Eines der gebräuchlichsten Verfahren ist die Vektorquantisierung mittels k-Means-Algorithmus. Allgemein teilt k-Means den Datensatz so in k-Partitionen auf, dass die Summe der quadratischen Abweichungen von den Zentren minimiert wird (hier mittels euklidischer Distanz). In vorliegender Untersuchung wurden die Startgruppen (VL bzw. keine VL) vorgegeben und zur Zentrumssuche der schnellere k-Means++-Algorithmus verwendet. Attribute mit gleichen Clusterzentren wurden eliminiert, um die Summe der quadratischen Fehler (SSE) zu minimieren [19,22,23].

Ergebnisse

Datensatz

1.451 Fälle von notärztlichen Atemwegssicherungen konnten aus mehr als 130.000 Datensätzen identifiziert werden. Hiervon war bei 459 Fällen der Atemweg unvollständig dokumentiert, sodass 922 Datensätze übrigblieben. Dabei wurde 277 Mal die Videolaryngoskopie verwendet (VL 28 %) und bei 715 darauf verzichtet (keine VL 72 %, Verhältnis im Originaldatensatz 1:2,6). Der Anteil männlicher Patienten in der Klasse VL betrug 73 %, in der Klasse ohne VL 75 % ($p = 0,34$). Bei den Attributen Erfahrung im Atemwegsmanagement fehlten insgesamt $n = 41$ (4 %), PES $n = 328$ (33 %), Herzfrequenz $n = 3$ (0,3 %), Atemfrequenz $n = 32$ (3 %) und Sauerstoffsättigung $n = 9$ (1 %) Angaben. Die Videolaryngoskopie wurde im Originaldatensatz zu 88,8 % primär verwendet, d. h. ohne Dokumentation eines Verfahrenswechsels im Atemwegsmanagement [8,9]. Eine supraglottische Atemwegshilfe wurde $n = 22$ (3,1 %) Mal ohne Angabe einer endotrachealen Intubation oder chirurgischen Atemwegssicherung benutzt.

Tabelle 1 gibt einen Überblick über die 26 Attribute und ihre Ausprägungen zwischen den beiden Gruppen. Zusammengefasst unterschieden sich die Gruppen in der Häufigkeit von Notärzten mit hohem Erfahrungsniveau (VL 85 % vs. keine VL 62 %, $p < 0,01$) und im Auskultationsbefund (Bronchiale Spastik

Tabelle 1

Vergleich der beiden Gruppen mit und ohne Videolaryngoskopie.

Attribut	Wert	VL (n = 277; 28 %)	Keine VL (n = 715; 72 %)	p-Wert
Erfahrung im Atemwegsmanagement	hoch mittel niedrig	85 % 6 % 9 %	62 % 17 % 21 %	<0,01* <0,01* <0,01*
Auskultation	Atemwegsverlegung/Apnoe/ Schnappatmung Bronchospasmus Rasselgeräusch Andere Normal	5 % 39 % 6 % 1 % 48 %	11 % 7 % 6 % 3 % 73 %	<0,01* <0,01* 0,6 0,1 <0,01*
Dyspnoe oder Zyanose		63 %	28 %	<0,01*
Kopfverletzung	keine/leicht mittel/schwer	33 %/6 % 20 %/41 %	34 %/4 % 14 %/48 %	0,8/0,2 0,02*/0,05*
Gesichtsverletzung	keine/leicht mittel/schwer	74 %/8 % 13 %/5 %	75 %/5 % 11 %/9 %	0,6/0,06 0,5/0,1
Verletzung HWS	keine/leicht mittel/schwer	93 %/1 % 4 %/1 %	88 %/3 % 5 %/4 %	0,02*/0,1 0,6/0,04*
Verletzung LWS/BWS	keine/leicht mittel/schwer	90 %/1 % 4 %/5 %	90 %/1 % 5 %/3 %	0,8/0,7 0,3/0,4
Thoraxtrauma	keine/leicht mittel/schwer	71 %/3 % 9 %/17 %	68 %/2 % 11 %/19 %	0,7/0,9 0,3/0,4
Abdominaltrauma	keine/leicht mittel/schwer	84 %/1 % 4 %/11 %	85 %/1 % 5 %/8 %	0,7/0,4 0,5/0,3
Beckenverletzung	keine/leicht mittel/schwer	84 %/0 % 4 %/12 %	83 %/2 % 5 %/10 %	0,7/0,05 0,3/0,3
Verletzung untere Extremität	keine/leicht mittel/schwer	77 %/2 % 7 %/14 %	75 %/4 % 9 %/13 %	0,4/0,1 0,4/0,7
Verletzung obere Extremität	keine/leicht mittel/schwer	77 %/4 % 14 %/5 %	79 %/6 % 8 %/7 %	0,4/0,4 <0,01*/0,3
Weichteilverletzung	keine/leicht mittel/schwer	91 %/4 % 3 %/2 %	88 %/3 % 5 %/4 %	0,2/0,4 0,2/0,1
Thoraxdrainage		14 %	12 %	0,6
Beckenschlinge		39 %	27 %	<0,01*
HWS-Immobilisation		79 %	67 %	<0,01*
Nichtinvasive Beatmung		70 %	19 %	<0,01*
Kapnographie		96 %	94 %	0,2
Schwierige Maskenbeatmung		1 %	0 %	0,3
Schwierige Atemwegsanatomie		6 %	4 %	0,2
> 2 Intubationsversuche		5 %	3 %	0,1
Muskelrelaxans		87 %	83 %	0,1
PES		2 (1 bis 2)	2 (1 bis 2)	0,8
Alter		54 [51,59–56,36]	56 [54,36–57,54]	0,2
GCS		15 (14 bis 15)	15 (14 bis 15)	0,8
Herzfrequenz	/min	90,3 [88,07–92,48]	89,2 [87,74–90,73]	0,4
Systolischer Blutdruck	mmHg	141 [137,6–144,5]	136 [133,76–138,24]	0,02*
Atemfrequenz	/min	16,5 [16,02–17,08]	15,8 [15,45–16,11]	0,02*
Sauerstoffsättigung	%	94,1 [93,36–94,74]	94,37 [93,83–94,9]	0,5

LWS: Lendenwirbelsäule; BWS: Brustwirbelsäule; HWS: Halswirbelsäule; GCS: Glasgow Coma Scale; Ergebnisse als Mittelwerte mit Standardabweichung oder Median mit Interquartilsabstand; alle Vitalwerte sind initiale Werte bei Erstkontakt; *statistisch signifikant Wert ($p < 0,05$).

VL 39 % vs. keine VL 7 %, Dyspnoe oder Zyanose VL 5 % vs. keine VL 11 %, beide $p < 0,01$). Mit Blick auf das Verletzungsmuster waren schwere Kopf- und Halswirbelverletzungen in der Klasse „Keine VL“ (41 % vs. 48 %, $p = 0,05$ und 1 % vs. 4 %, $p = 0,04$) häufiger, wohingegen mittlere Verletzungen der unteren Extremität hier seltener waren (14 % vs. 8 %, $p < 0,01$). Eine Immobilisation der Halswirbelsäule (HWS) wurde signifikant häufiger in der Klasse „VL“ durchgeführt (79 % vs. 67 %, $p < 0,01$), ebenso wie die nichtinvasive Beatmung (NIV) (70 % vs. 19 %, $p < 0,01$). Bei den initialen Vitalwerten waren der systolische Blutdruck sowie die Atemfrequenz signifikant höher in der Klasse „VL“ (141 mmHg [137,6–144,5] vs. 136 mmHg [133,76–138,24] und 16,5/min [16,02–17,08] vs. 15,8/min [15,45–16,11], beide $p = 0,02$).

Diskriminationsfähigkeit bei steigender Verwendung

Wie in Tabelle 2 aufgeführt konnte im Originaldatensatz eine Gesamtkorrektheit von 79,05 % [95%-Konfidenzin-

tervall 78,77–79,34] erzielt werden bei einem Kappa von 0,44 [0,43–0,45]. Die Sensitivität für die Anwendung der VL lag bei 0,66 [0,66–0,67] mit einem PPV von 0,5 [0,5–0,51]. Die Spezifität lag bei 0,82 [0,82–0,83] bei einem NPV von 0,9 [0,9–0,9]. Falsch negative bzw. positive Rate lagen bei 0,34 [0,33–0,34] bzw. 0,18 [0,17–0,18]. Die AUC-ROC betrug 0,78 [0,78–0,79].

Im simulierten Verhältnis 1:1 der VL erreichte das Modell eine Sensitivität von 0,77 [0,76–0,79] bei einem PPV von 0,66 [0,66–0,67]. Die Spezifität verminderte sich auf 0,71 [0,70–0,71] bei einem NPV von 0,8 [0,79–0,82]. Die AUC-ROC betrug 0,78 [0,77–0,79] bei jedoch fallender Gesamtkorrektheit (73,39 % [72,41–74,38]) und fallendem Kappa 0,47 [0,45–0,49]. Die falsch negativen bzw. positiven Raten lagen bei 0,23 [0,21–0,24] bzw. 0,29 [0,29–0,3] (alle $p < 0,01$).

Mit simulierter steigender Verwendung der VL (Verhältnis 2:1) steigerte sich die Sensitivität auf 0,77 [0,76–0,78] bei einem PPV von 0,82 [0,83–0,87]. Die Spezifität fiel auf 0,58 [0,58–0,59] bei

einem NPV von 0,51 [0,48–0,54] ab. Die AUC-ROC betrug 0,77 [0,75–0,78] bei jedoch fallender Gesamtkorrektheit (71,66 % [71–72,33]) und fallendem Kappa 0,34 [0,32–0,36]. Falsch negative bzw. positive Rate lagen bei 0,23 [0,22–0,24] bzw. 0,42 [0,41–0,42] (alle $p < 0,01$ zum Original bzw. 1:1). AUC-ROC und Precision-Recall-Kurven sind in Abbildung 2 dargestellt.

Attributgewichtung

Innerhalb des Random-Forest-Modells war das Patientenalter (0,447) das am höchsten gewichtete Attribut gefolgt vom initialen Blutdruck (0,376), Kopfverletzungen (0,369) sowie vom Erfahrungsgrad des Notarztes (0,351). Verglichen zum Originalverhältnis war bis zum Simulationsverhältnis 2:1 ein starker Verlust beim Erfahrungsgrad (-0,045) sowie ein starker Zugewinn beim Attribut schwierige Maskenbeatmung (+0,069) zu ermitteln (Tab. 3 und Abb. 3).

Clustering

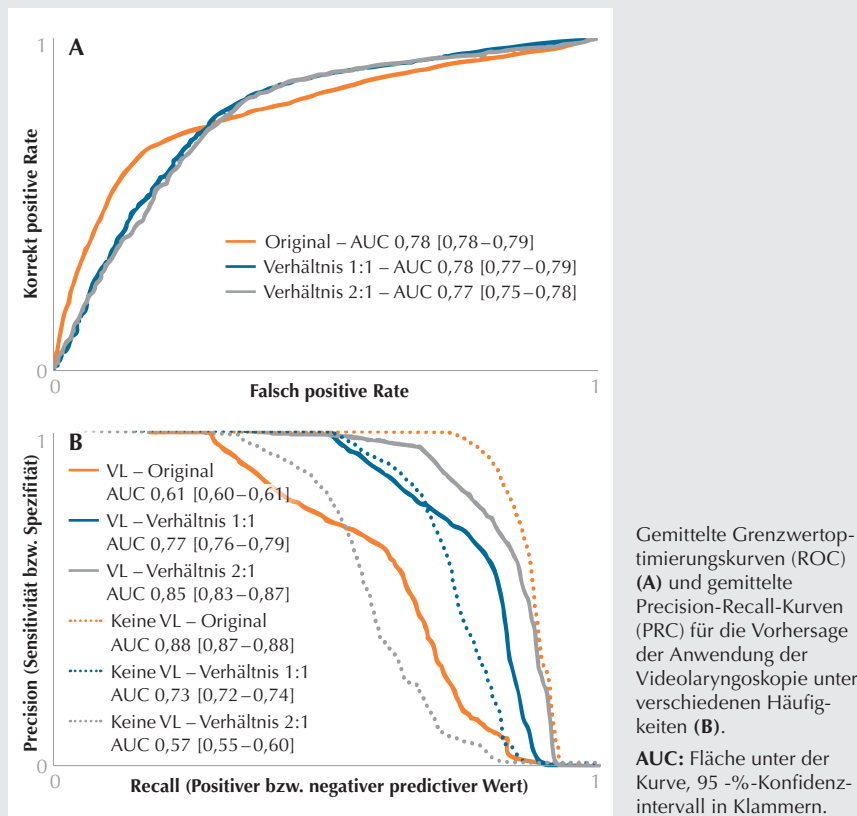
Die Clusteranalyse im Originaldatensatz mittels k-Means-Algorithmus ergab den

Tabelle 2 Vergleich der Resultate des Random-Forest-Algorithmus zur Vorhersage der Verwendung der Videolaryngoskopie beim schwerverletzten Patienten unter verschiedenen Anwendungshäufigkeiten

Gütekriterium	Original 1:2,6	Verhältnis 1:1	Verhältnis 2:1	p-Wert Original vs. 1:1 bzw. 1:1 vs. 2:1
Gesamtkorrektheit (%)	79,05 [78,77–79,34]	73,39 [72,41–74,38]	71,66 [71–72,33]	Beide $< 0,01^*$
Kappa	0,44 [0,43–0,45]	0,47 [0,45–0,49]	0,34 [0,32–0,36]	Beide $< 0,01^*$
AUC-ROC	0,78 [0,78–0,79]	0,78 [0,77–0,79]	0,77 [0,75–0,78]	Beide $< 0,01^*$
MCC	0,45 [0,44–0,45]	0,47 [0,45–0,49]	0,34 [0,32–0,36]	Beide $< 0,01^*$
Sensitivität (VL)	0,66 [0,66–0,67]	0,77 [0,76–0,79]	0,77 [0,76–0,78]	Beide $< 0,01^*$
Positiv prädiktiver Wert	0,5 [0,5–0,51]	0,66 [0,66–0,67]	0,82 [0,81–0,83]	Beide $< 0,01^*$
Falsch positiver Wert	0,18 [0,17–0,18]	0,29 [0,29–0,3]	0,42 [0,41–0,42]	Beide $< 0,01^*$
AUC-PRC	0,61 [0,6–0,61]	0,77 [0,76–0,79]	0,85 [0,83–0,87]	Beide $< 0,01^*$
Spezifität (Keine VL)	0,82 [0,82–0,83]	0,71 [0,7–0,71]	0,58 [0,58–0,59]	Beide $< 0,01^*$
Negativ prädiktiver Wert	0,9 [0,9–0,9]	0,8 [0,79–0,82]	0,51 [0,48–0,54]	Beide $< 0,01^*$
Falsch negativer Wert	0,34 [0,33–0,34]	0,23 [0,21–0,24]	0,23 [0,22–0,24]	Beide $< 0,01^*$
AUC-PRC	0,88 [0,87–0,88]	0,73 [0,72–0,74]	0,57 [0,55–0,6]	Beide $< 0,01^*$
Schwellenwert	0,5 [0,5–0,5]	0,5 [0,5–0,5]	0,5 [0,5–0,5]	Nicht untersucht

95%-Konfidenzintervall in Klammern, das Originalverhältnis war VL/keine VL 1:2,6.
AUC-ROC: Fläche unter Grenzwertoptimierungskurve; MCC: Matthews Korrelationskoeffizient; AUC-PRC: Fläche unter der Precision-Recall-Kurve (Sensitivität zu positiv prädiktivem Wert bzw. Spezifität zu negativ prädiktivem Wert); * statistisch signifikant Wert ($p < 0,05$).

Abbildung 2



Cluster „keine VL“ (n = 744) mit den Attributen unauffällige Auskultation, keine Dyspnoe oder Zyanose, keine NIV, Patientenalter = 57,3, GCS = 13,4, Herzfrequenz = 88,7/min, systolischer Blutdruck = 137,6 mmHg, Atemfrequenz = 15,7/min, Sauerstoffsättigung = 94 %. Der Cluster „VL“ (n = 248) enthielt die Attribute Spastik, Dyspnoe oder Zyanose, NIV, Patientenalter = 49,7, GCS = 12,4, Herzfrequenz = 92/min, systolischer Blutdruck = 136 mmHg, Atemfrequenz = 16,7/min, Sauerstoffsättigung = 92,9 %. Im Clustering wurden so 78,53 % korrekt klassifiziert bei einer Summe der quadratischen Fehler von 623,9.

Diskussion

Die vorliegende Studie zeigte, dass trotz der Nutzung von Echtweltdaten das im Originaldatensatz entwickelte Random-Forest-Modell eine bereits akzeptable Diskriminationsfähigkeit (AUC-ROC

0,77) aufweist und diese in allen Verhältnissimulationen konstant blieb. Besonders gewichtige Attribute im Modell waren u. a. das Patientenalter, Vitalwerte und Kopfverletzungen sowie im Clustering Spastik, NIV, Dyspnoe oder Zyanose, Alter, GCS und Vitalwerte. Unabhängig von den limitierenden Faktoren in den Echtweltdaten zeigte sich zudem in den Modellierungen zur steigenden Verwendungshäufigkeit eine Erhöhung des PPV bei gleichzeitiger Abnahme des NPV. Die Untersuchung ist die erste, welche die Anwendungshäufigkeit der Videolaryngoskopie mittels maschinellen Lernens analysiert und zeigt, dass prinzipiell eine Modellierung der Anwendung der VL trotz der Limitierungen durch Echtweltdaten bzw. die Datenstruktur möglich erscheint. Somit ist ein Proof of Concept gegeben, welches im von Fleuren et al. vorgeschlagenen und an die Raumfahrt adaptieren Entwicklungsprozess des Maschinellen Lernens die Stufe 3 von 10 (Modellprototyp und

-entwicklung) einnehmen würde [24]. Welche Konsequenzen sich für die Leitlinien-treue sowie für eine Erweiterung der Datenstruktur und zukünftige Modelle ableiten lassen, soll nun im Folgenden diskutiert werden.

Implikationen für die Leitlinien-treue

Menschen arbeiten mit den Werkzeugen, mit denen sie den höchsten Erfahrungsschatz besitzen [18]. Daher stellt der Kerngedanke der aktuell gültigen Leitlinie zum prähospitalen Atemwegsmanagement, die VL für die endotracheale Intubation konventionellen Laryngoskopen wegen der nachgewiesenen höheren Rate an primär erfolgreichen Intubationen vorzuziehen, einen Paradigmenwechsel dar [2,6]. Bei weiterbestehen der präklinischen Vorhaltung von konventionellen Laryngoskopen ist leider aber eine nahezu dauerhafte Verwendung der VL als unwahrscheinlich anzusehen. Die hier vorliegenden Modelle könnten aber den Kerngedanken der Leitlinie unterstützen: Bei niedriger Anwendungshäufigkeit wäre aufgrund hoher Spezifitäts- bzw. NPV-Werte eine Hinweisfunktion zur Abweichung von der Leitlinie denkbar. Bei gesteigerter Anwendungshäufigkeit könnte genau der umgekehrte Effekt auftreten, da hier die Gewichtung hin zum Einsatz der VL den Nutzer in seinem Vorgehen bestärken könnte. Zudem würde nur ein falsch positiver Wert, d. h. die Vorhersage der VL bei tatsächlichem Verzicht hierauf, zu einer Schwächung der Leitlinie führen ($p = 0,18$ im Originaldatensatz bzw. $p = 0,42$ beim Verhältnis 2:1). Eine falsch negative Vorhersage, d. h. die tatsächliche VL-Anwendung bei Vorhersage Verzicht ($p = 0,34$ im Originaldatensatz bzw. $p = 0,23$ beim Verhältnis 2:1), würde klinisch nicht ins Gewicht fallen. Somit könnte die vorliegende, aber auch zukünftige Modellierungen durchaus die Ziele der aktuellen Leitlinie unterstützen. Hierfür bedarf es jedoch einer Weiterentwicklung, um eine stärkere Robustheit zu gewährleisten. Auch das überwachte Clustering ergab Hinweise auf Anwendungscharakteristika. Denkbar wäre, diese Ergebnisse weiter aufzu-

Tabelle 3
Vergleich der Attributgewichtungen innerhalb des Random-Forest-Algorithmus zur Vorhersage der Verwendung der Videolaryngoskopie beim schwerverletzten Patienten unter verschiedenen Anwendungshäufigkeiten.

Attribut	Gewicht Original	Gewichtung Verhältnis 1:1	Gewichtung Verhältnis 2:1	p-Wert		
				Original vs. 1:1	Original vs. 2:1	1:1 vs. 2:1
> 2 Intubationsversuche	0,27 [0,26–0,27]	0,29 [0,2–0,31]	0,3 [0,28–0,32]	0,02*	0,01*	0,37
Abdominelle Verletzungen	0,28 [0,27–0,29]	0,3 [0,29–0,31]	0,3 [0,29–0,32]	0,01*	0,02*	0,74
Patientenalter	0,45 [0,44–0,45]	0,45 [0,45–0,46]	0,45 [0,44–0,46]	0,06	0,19	1,00
Atemfrequenz initial	0,3 [0,29–0,3]	0,32 [0,32–0,32]	0,32 [0,31–0,33]	<0,01*	<0,01*	0,68
Auskultationsbefund	0,31 [0,3–0,32]	0,31 [0,3–0,32]	0,32 [0,31–0,33]	0,50	0,07	0,23
Dyspnoe oder Zyanose	0,34 [0,33–0,35]	0,36 [0,34–0,37]	0,36 [0,35–0,38]	0,02*	< 0,01*	0,35
Erfahrung im Atemwegsmanagement	0,35 [0,34–0,36]	0,38 [0,37–0,38]	0,4 [0,39–0,4]	<0,01*	<0,01*	<0,01*
GCS	0,33 [0,33–0,34]	0,34 [0,34–0,35]	0,34 [0,33–0,35]	0,01*	0,03*	0,83
Gesichtsverletzung	0,31 [0,3–0,31]	0,32 [0,31–0,33]	0,32 [0,31–0,33]	<0,01*	0,03*	0,77
Herzfrequenz initial	0,34 [0,34–0,35]	0,36 [0,35–0,37]	0,37 [0,36–0,37]	<0,01*	<0,01*	0,03*
Halswirbelsäulenverletzung	0,25 [0,25–0,26]	0,27 [0,26–0,28]	0,28 [0,27–0,3]	0,01*	<0,01*	0,20
Kapnographie	0,29 [0,28–0,3]	0,31 [0,29–0,33]	0,32 [0,3–0,33]	0,05*	0,02*	0,56
Kopfverletzung	0,37 [0,36–0,38]	0,37 [0,36–0,38]	0,38 [0,37–0,39]	0,70	0,05	0,04*
Verletzung der restlichen Wirbelsäule	0,26 [0,26–0,27]	0,28 [0,27–0,29]	0,28 [0,27–0,29]	<0,01*	<0,01*	1,00
Muskelrelaxans	0,31 [0,31–0,32]	0,34 [0,32–0,35]	0,33 [0,31–0,34]	0,01	0,06	0,24
Nichtinvasive Beatmung	0,28 [0,27–0,29]	0,29 [0,28–0,3]	0,29 [0,28–0,3]	0,10	0,22	0,58
Pre-Emergency-Status	0,31 [0,31–0,32]	0,32 [0,31–0,33]	0,32 [0,31–0,33]	0,13	0,39	0,88
Systolischer Blutdruck	0,38 [0,37–0,38]	0,38 [0,37–0,39]	0,37 [0,36–0,38]	0,30	0,67	0,28
Sauerstoffsättigung	0,31 [0,31–0,32]	0,33 [0,33–0,34]	0,33 [0,32–0,33]	<0,01*	< 0,01*	0,37
Schwierige Atemwegsanatomie	0,3 [0,29–0,32]	0,31 [0,29–0,33]	0,31 [0,29–0,32]	0,72	0,69	1,00
Schwierige Maskenbeatmung	0,26 [0,24–0,29]	0,27 [0,22–0,31]	0,19 [0,12–0,26]	0,77	0,06*	0,06*
Supraglottische Atemwegshilfe	0,31 [0,29–0,32]	0,31 [0,29–0,32]	0,3 [0,28–0,32]	0,67	0,70	0,51
Immobilisation der Halswirbelsäule	0,34 [0,33–0,34]	0,35 [0,34–0,36]	0,36 [0,34–0,37]	0,03*	0,02*	0,50
Thoraxverletzung	0,33 [0,32–0,34]	0,34 [0,33–0,35]	0,34 [0,33–0,35]	0,03*	0,17	0,74
Thoraxdrainage	0,31 [0,3–0,32]	0,31 [0,29–0,33]	0,31 [0,3–0,33]	0,85	1,00	0,86
Weichteilverletzung	0,26 [0,25–0,26]	0,26 [0,25–0,27]	0,27 [0,25–0,28]	0,16	0,20	0,80

95-%-Konfidenzintervall in Klammern; **GCS**: Glasgow Coma Scale; das Originalverhältnis war VL/Keine VL 1:2,6; alle Vitalwerte sind initiale Werte bei Erstkontakt; *statistisch signifikant Wert ($p < 0,05$).

greifen und in einem größeren Umfeld zu untersuchen, um später den Einsatz der VL für bestimmte Situationen, in denen keine Anwendung routinemäßig stattfindet, weiter zu fördern.

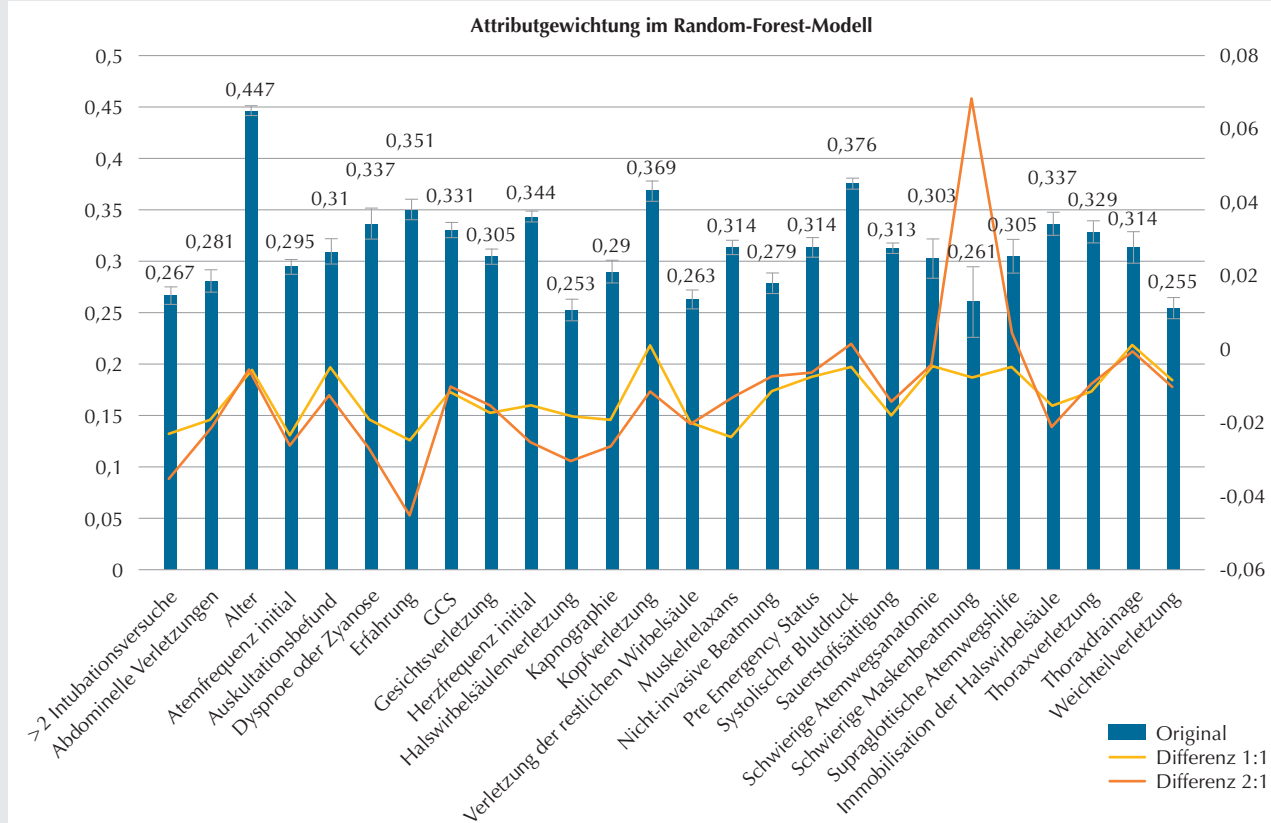
Implikationen für die Modellentwicklung

Koch et al. konnten in ihrer aktuellen Umfrage zeigen, dass zum einen die Einführung der Leitlinie zum prähospitalen Atemwegsmanagement zu einem Umdenken hin zur primären Verwendung der VL bei 39 % der Notärzte geführt

hat. Andererseits gaben immer noch 55 % eine sekundäre prähospitaler Verwendung an. In ihrer Regressionsanalyse identifizierten sie die Attribute Notärztin, VL C-MAC® oder McGrath® und Leitlinienkenntnisse als mit dem primären Einsatz der VL assoziiert. Fachrichtung, Transportmodalität, Alter des Notarztes und Erfahrung im Atemwegsmanagement korrelierten in ihrer Untersuchung dagegen nicht. Als häufigsten Grund auf Verzicht nannten sie fehlende Erfahrung (7 %), keine Verfügbarkeit (6 %) bzw. dass die VL keinen Vorteil in der Atem-

wegssicherung bieten würde (13 %) [7]. Andere, in unserer Studie identifizierte und bei Koch et al. nicht erfasste Attribute waren das Patientenalter, Kopf-Hals-Verletzungen oder Bronchospastik, welche alle auf die Hypoxietoleranz, alternative Atemwegshilfsmittel, die Atemwegsanatomie bzw. die Sichtachse (z. B. durch entfernbare Gebissprothesen) maßgeblichen Einfluss haben können [25]. Als Kontraindikationen für den Einsatz der VL gelten u. a. die massiv eingeschränkte Mundöffnung, welche die Platzierung eines Spatels behindert

Abbildung 3



Veränderung der Gewichtung der Attribute innerhalb des Random-Forrest-Algorithmus unter Berücksichtigung der simulierten Anwendungshäufigkeit.

[5,26]. So wird beispielsweise eine Kombination aus verminderter Mundöffnung und eingeschränkter Halswirbelsäulenmobilisation bei notfallmäßiger Indikation zur invasiven Atemwegssicherung, welche leitlinienkonform die primäre Anwendung supraglottischer Atemwegshilfen rechtfertigen würde, im MIND nicht abgebildet, ebenso wenig, ob ein (und falls ja, welcher) Gerätetyp vorgehalten wird [27].

Aus den Ergebnissen dieses Proof of Concept sowie den Studienergebnissen von Koch et al. lässt sich jedoch skizzieren, wie eine computergestützte Unterstützung in der prähospitalen Atemwegssicherung aussehen könnte und wie Anwender, die der VL eher skeptisch gegenüberstehen, erreicht werden könnten. Ein denkbarer Weg wäre, zukünftig beispielsweise die Effektivität gewählter Atemwegssicherungsmaßnahmen

patienten- und anwenderindividuell zu modellieren. Hierzu müsste vorab geklärt werden, welche Anwender- und Patientengruppen von einem Verzicht oder der Anwendung der VL profitieren bzw. bei welchen die Maßnahmen ohne Gefährdung einhergehen würden. Ein solches Modell sollte zum einen aus einem übergeordneten Nutzer- und Standortprofil, welches das individuelle Ausbildungsniveau des Notarztes sowie die vorgehaltenen und verfügbaren Atemwegshilfsmittel beinhaltet, bestehen. Neben Stichworten zur Indikation der Atemwegssicherung könnten die Umgebungsbedingungen, die Atemwegs-anatomie und zumindest die ersten drei Versuche analog zum Leitlinienalgorithmus detaillierter registriert werden. Abgesehen von Prädiktoren für schwierige Atemwege könnten das gewählte Mittel sowie gegebenenfalls der Abbruchgrund miterfasst werden. Da das

Atemwegsmanagement und gerade die Verwendung der Videolaryngoskopie Qualitätsmerkmale der präklinischen Versorgung von Notfallpatienten sind, wäre aus Sicht der Autoren eine Diskussion über ein vergleichbares Register wie dem in der Reanimatologie implementiertem sinnvoll [28]. Jedoch müssten vorab datenschutzrechtliche Bedenken geklärt und eine ausufernde Dokumentation vermieden werden. Im sich neu definierenden Arzt-Computer-Verhältnis muss allen Beteiligten die Sinnhaftigkeit und Effektivität solch eines Systems, welches nicht nur rein wissenschaftlichen Zwecken, sondern auch der Patientenversorgung dienen soll, bewusst sein, um einer mangelnden Nutzerakzeptanz respektive einer verminderten Datenqualität vorzubeugen. Als Alternative wäre allerdings auch denkbar, die direkte Zugänglichkeit zur VL durch organisatorische Maßnahmen

zu steigern, z. B. durch den totalen Ersatz der konventionellen Laryngoskope oder lediglich die Vorhaltung in einem Ersatzrucksack. Dieser Paradigmenwechsel müsste dann allerdings mit Blick auf die notärztliche Ausbildung auch innerklinisch stattfinden und die Atemwegssicherung bei allen (Notfall-)Patienten primär mittels Videolaryngoskopie etabliert werden, wie es Cook et al. oder De Jong et al. in ihren Untersuchungen bereits durchgeführt haben [6,29–31].

Limitationen

Es handelt sich bei vorliegender Studie um eine Sekundärauswertung von präklinischen Polytraumadaten aus einer deutschen Region, sodass diese einer Überprüfung an einer unabhängigen Kohorte bedarf. Die Ergebnisse sind nicht auf internistische, neurologische oder pädiatrische Patienten übertragbar. Die Leitlinie zur präklinischen Atemwegshilfe wurde im Untersuchungszeitraum veröffentlicht, weswegen ihre Effekte auf die Gerätevorhaltung und Training eventuell noch nicht in der notärztlichen Versorgung spürbar waren. Jedoch enthielt die vorherige Handlungsempfehlung bereits eine situativ positive Bewertung der Videolaryngoskopie [2,32]. Gerade mit Blick auf die COVID-Pandemie und die Empfehlungen zur primären Verwendung der VL bei der invasiven Atemwegssicherung dieser Patienten könnte es zu einer deutlichen, fächerübergreifenden Verbreitung der Technik gekommen sein [33]. Da in vorliegender Studie die Evaluation des MIND zur Vorhersage der Verwendung der VL im Vordergrund stand, wurde auf eine Testung weiterer Algorithmen des maschinellen Lernens verzichtet. Einschränkend muss mit Blick auf die zur Verfügung stehenden Attribute nochmals betont werden, dass der MIND leider nur die Vitalwerte bei notärztlichem Erstkontakt und Übergabe erfasst und somit eine Erststabilisierung durch Notfallsanitäter oder Veränderungen nicht mitefassen werden. Somit kann schlussendlich auch keine Aussage zur korrekten Indikationsstellung für eine Notfallnarkose getroffen werden. Die Selektion auf Narkosen mit dokumentiertem Atemwegsmanagement

stellt qualitativ optimierte Echtweltdaten mit Blick auf die notärztliche Dokumentationsgenauigkeit dar. Der Verlust von gut 31 % ist in der hiesigen Grundlagenarbeit nur von geringem Effekt und käme erst bei einem finalen Modell zu tragen. Verglichen zu vorherigen Untersuchungen ergab sich so eine deutliche Steigerung der Verwendungsrate der VL von 4,6 % (2015–2017) auf 28 % (2018–2020) [8]. Generell eine größere prähospital Akzeptanz hieraus abzuleiten ist jedoch bei fehlender Datenlage zu Verfügbarkeit, Anschaffungszeitraum respektive Modellvorhaltung nicht sicher möglich. Im Originaldatensatz zeigte sich weder eine Klassenimbalance noch ein Minoritätsproblem [34,35]. Der Datensatz erschien zudem für weitere Aufteilungen (z. B. in Training und Testdatensatz) zu gering, weswegen dahingehende Strategien verworfen wurden. Zusätzlich bestehende internistische oder neurologische Krankheitsbilder spielten im ursprünglichen Datensatz keine Rolle.

Schlussfolgerung

Zusammenfassend konnte in vorliegender Studie als Proof of Concept eine stabile Diskriminationsfähigkeit durch eine konstante AUC-ROC bei verschiedenen Anwendungshäufigkeiten der Videolaryngoskopie für polytraumatisierte Patienten auf Echtweltdaten des MIND gezeigt werden. Ebenso wurden unterschiedliche Verwendungcluster identifiziert. Prinzipiell könnten solche Modelle den Kerngedanken der Leitlinie, nämlich die bevorzugte Verwendung der VL, unterstützen. Dafür müsste die Robustheit der Attribute jedoch durch eine Datensatzerweiterung erhöht werden. Bis ein konsensbasiertes, computergestütztes Unterstützungstool zum Atemwegsmanagement auf Grundlage eines erweiterten Datensatzes hieraus entwickelt und validiert worden ist, könnten auch organisatorische Maßnahmen getroffen werden, um die Anwendungshäufigkeit der Videolaryngoskopie zu steigern.

Literatur

1. Bernhard M, Matthes G, Kanz KG, Waydhas C, Fischbacher M, Fischer M, et al: Notfallnarkose, Atemwegsmanagement und Beatmung beim Polytrauma: Hintergrund und Kernaussagen der interdisziplinären S3-Leitlinie Polytrauma. *Anaesthesist* 2011;60:1027–1040
2. Timmermann A, Böttiger BW, Byhahn C, Dörge V, Eich C, Gräsner J-T, et al: S1-Leitlinie Prähospitaler Atemwegsmanagement (Kurzfassung). *Anaesth Intensivmed* 2019;6:316–336
3. Deutsche Gesellschaft für Unfallchirurgie, et al: S3-Leitlinie Polytrauma/Schwerer Verletzten-Behandlung (AWMF Registernummer 187-023), Version 4.0 (31.12.2022); 2023. <https://www.awmf.org/leitlinien/detail/II/187-023.html> (Zugriffsdatum: 08.02.2024)
4. Hossfeld B, Frey K, Doerges V, Lampl L, Helm M: Improvement in glottic visualisation by using the C-MAC PM video laryngoscope as a first-line device for out-of-hospital emergency tracheal intubation: An observational study. *Eur J Anaesthesiol* 2015;32:425–431
5. Lewis SR, Butler AR, Parker J, Cook TM, Smith AF: Videolaryngoscopy versus direct laryngoscopy for adult patients requiring tracheal intubation. *Cochrane Database Syst Rev* 2016. DOI: 10.1002/14651858.CD011136.pub2
6. Hansel J, Rogers AM, Lewis SR, Cook TM, Smith AF: Videolaryngoscopy versus direct laryngoscopy for adults undergoing tracheal intubation. *Br J Anaesth*;129:612–623
7. Koch DA, Hagebusch P, Faul P, Steinfeldt T, Hoffmann R, Schweigkofler U: Analysis of the primary utilization of videolaryngoscopy in prehospital emergency care in Germany. *Anaesthesiologie* 2023;72:245–252
8. Luckscheiter A, Lohs T, Fischer M, Zink W: Notärztliche Durchführung von Narkosen: Eine Istanalyse der Jahre 2015–2017. *Anaesthesist* 2019;68:270–281
9. Luckscheiter A, Lohs T, Fischer M, Zink W: Notärztliches Atemwegsmanagement in Abhängigkeit von Fachrichtung und Weiterbildung. *Anaesthesist* 2020;69:170–182
10. Landucci F, Caldiroli D, Chiumello D, Byrne A: Cognitive resistance towards videolaryngoscopy... and why Macintosh refuses to die. *Minerva Anestesiol* 2017;83:1221–1223
11. TraumaRegister DGU Standardbogen V2020 (06/20). <https://www.auc-online.de>

- de/fileadmin/AUC/Dokumente/Register/TraumaRegister_DGU/TR-DGU_-_Standardbogen.pdf (Zugriffsdatum: 08.02.2024)
12. Reanimationsdatensatz EV 2.1.1. <https://www.reanimationsregister.de/downloads/datensatzdefinitionen/51-datensatz-ev-2-1-1/file.html> (Zugriffsdatum: 08.02.2024)
 13. Messelken M, Schlechtriemen T, Arntz H-R, Bohn A, Bradschettl G, Brammen D, et al: Der Minimaler Notfalldatensatz MIND3. *Notfall Rettungsmed* 2011;14:647–654
 14. Luckscheiter A, Zink W, Lohs T, Eisenberger J, Thiel M, Viergutz T: Machine learning for the prediction of preclinical airway management in injured patients: a registry-based trial. *Clin Exp Emerg Med* 2022;9:304–313
 15. Liu F, Panagiotakos D: Real-world data: a brief review of the methods, applications, challenges and opportunities. *BMC Med Res Methodol* 2022;22:287
 16. Collins GS, Reitsma JB, Altman DG, Moons KGM: Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMC* 2015;13:1
 17. Bouillon B, Pieper D, Flohé S, Eikermann M, Prengel P, Ruchholtz S, et al: Level 3 guideline on the treatment of patients with severe/multiple injuries. *European J Trauma Emerg Surg* 2018;44:3–271
 18. Breckwoldt J, Klemstein S, Brunne B, Schnitzer L, Arntz H-R, Mochmann H-C: Expertise in prehospital endotracheal intubation by emergency medicine physicians – Comparing ‘proficient performers’ and ‘experts.’ *Resuscitation* 2012;83:434–439
 19. Witten IH, Frank E, Hall MA, Pal CJ: Data Mining. *Practical Machine Learning Tools and Techniques* (Fourth Edition). Morgan Kaufmann 2017
 20. Nahm FS: Receiver operating characteristic curve: overview and practical use for clinicians. *Korean J Anesthesiol* 2022;75:25–36
 21. Chicco D, Tötsch N, Jurman G: The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Min* 2021;14:13
 22. Drab K, Daszykowski M: Clustering in Analytical Chemistry. *J AOAC Int* 2014;97:29–38
 23. Peine A, Lütge C, Poszler F, Celi LA, Schöffski O, Marx G, et al: Künstliche Intelligenz und maschinelles Lernen in der intensivmedizinischen Forschung und klinischen Anwendung. *Anästh Intensivmed* 2020;372–384
 24. Fleuren LM, Thorat P, Shillan D, Ercole A, Elbers PWG: Machine learning in intensive care medicine: ready for take-off? *Intensive Care Med* 2020;46:1486–1488
 25. Breckwoldt J, Klemstein S, Brunne B, Schnitzer L, Mochmann H-C, Arntz H-R: Difficult prehospital endotracheal intubation – predisposing factors in a physician based EMS. *Resuscitation* 2011;82:1519–1524
 26. Hypes C, Sakles J, Joshi R, Greenberg J, Natt B, Malo J, et al: Failure to achieve first attempt success at intubation using video laryngoscopy is associated with increased complications. *Intern Emerg Med* 2017;12:1235–1243
 27. Lee J, Cho Y, Kim W, Choi K-S, Jang B-H, Shin H, et al: Comparisons of Videolaryngoscopes for Intubation Undergoing General Anesthesia: Systematic Review and Network Meta-Analysis of Randomized Controlled Trials. *J Pers Med* 2022;12:363
 28. Sunde GA, Kottmann A, Heltne JK, Sandberg M, Gellerfors M, Krüger A, et al: Standardised data reporting from pre-hospital advanced airway management – a nominal group technique update of the Utstein-style airway template. *Scand J Trauma Resusc Emerg Med* 2018;26:46
 29. Cook TM, Boniface NJ, Seller C, Hughes J, Damen C, MacDonald L, et al: Universal videolaryngoscopy: a structured approach to conversion to videolaryngoscopy for all intubations in an anaesthetic and intensive care department. *Br J Anaesth* 2018;120:173–180
 30. De Jong A, Sfara T, Pouzeratte Y, Pensier J, Rolle A, Chanques G, et al: Videolaryngoscopy as a first-intention technique for tracheal intubation in unselected surgical patients: a before and after observational study. *Br J Anaesth* 2022;129:624–634
 31. Piepho T, Cavus E, Noppens R, Byhahn C, Dörger V, Zwissler B, et al: S1-Leitlinie Atemwegsmanagement. *Anaesthesist* 2015;64:859–873
 32. Timmermann A, Byhahn C, Wenzel V, Eich C, Piepho T, Bernhard M, et al: Handlungsempfehlung für das präklinische Atemwegsmanagement. *Anästh Intensivmed* 2012;53:294–308
 33. Schälte G, Kehl F, Didion N, Pirlich N, Bickenbach J, Rehm M, et al: Besonderheiten des Atemwegsmanagements bei Patienten mit vermuteter oder gesicherter COVID-19 Erkrankung und bei Patienten ohne Infektion während der Corona-Pandemie. *Anästh Intensivmed* 2020;61:132–136
 34. Ellis RJ, Sander RM, Limon A: Twelve key challenges in medical machine learning and solutions. *Intelligence-Based Medicine* 2022;100068. <https://doi.org/10.1016/j.ibmed.2022.100068>
 35. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research* 2002;16:321–357.

Korrespondenz- adresse

Dr. med.
André Luckscheiter



Klinik für Anästhesiologie, Operative
Intensivmedizin und Notfallmedizin
Klinikum Ludwigshafen
Bremserstraße 79
67063 Ludwigshafen am Rhein,
Deutschland

Tel.: 0621 503-3000

Fax: 0621 503-30024

E-Mail: luckscha@klilu.de

ORCID-ID: 0000-0002-5724-7130