

## Tomorrow is already here [1] How machine learning is influencing anaesthesiology and intensive care medicine

S. Kagerbauer<sup>1</sup> · M. Blobner<sup>1</sup> · B. Ulm<sup>1</sup> · B. Jungwirth<sup>1,2</sup>



www.ai-online.info

► **Zitierweise:** Kagerbauer S, Blobner M, Ulm B, Jungwirth B: Die Zukunft hat schon begonnen. Wie maschinelles Lernen Anästhesie und Intensivmedizin prägt. *Anästh Intensivmed* 2020;61:85–96. DOI: 10.19224/ai2020.85

- 1 Klinik für Anästhesiologie und Intensivmedizin, Technische Universität München, (Direktor: Prof. Dr. G. Schneider)
- 2 Klinik für Anästhesiologie, Universitätsklinikum Ulm (Ärztliche Direktorin: Prof. Dr. B. Jungwirth)

### Interessenkonflikt

Die Autoren geben an, dass keine Interessenkonflikte bestehen.

### Schlüsselwörter

Künstliche Intelligenz –  
Maschinelles Lernen –  
Big Data – Risikoprädiktion

### Keywords

Artificial Intelligence –  
Machine Learning – Big Data –  
Risk Prediction

### Zusammenfassung

Künstliche Intelligenz ist aus der modernen Medizin nicht mehr wegzudenken. Gerade in Anästhesie und Intensivmedizin, wo häufig elektronische Patientendatenmanagementsysteme eingesetzt werden, führen wachsende Speicherkapazitäten und neue Möglichkeiten der Datenverarbeitung dazu, dass immer größere Datenmengen erhoben und analysiert werden. Um maschinelles Lernen darauf anwenden zu können, müssen Daten nicht nur einfach gespeichert, sondern auch leicht zugänglich, von verschiedenen Systemen zu bearbeiten und wiederholt zu benutzen sein. Als Analysemethoden werden verschiedenste Verfahren des überwachten und unüberwachten Lernens angewendet. Auf diese Art und Weise erfolgen unter anderem die Hypothesengenerierung für randomisierte prospektive Studien, die Analyse seltener Komplikationen, die Risikoberechnung oder die Entwicklung von Entscheidungsunterstützungssystemen. Dieser Artikel soll eine Übersicht über Methoden und Anwendungsbereiche maschinellen Lernens in der Anästhesiologie und Intensivmedizin geben. Darüber hinaus soll aufgezeigt werden, welche Probleme mit den neuen Technologien verbunden sein können und wie man sie lösen kann. Ziel der Anwendung von maschinellem Lernen in der Anästhesiologie und Intensivmedizin wie auch in anderen Fachgebieten der Medizin ist, durch eine personalisierte Medizin Komplikationen zu vermeiden und die Behandlungsqualität zu steigern.

## Die Zukunft hat schon begonnen [1] Wie maschinelles Lernen Anästhesie und Intensiv- medizin prägt

### Summary

Artificial intelligence has become an everyday part of modern medicine. Increasing storage capacity and new ways of processing data are leading to ever increasing quantities of data being collected and analysed, especially in the areas of anaesthesia and intensive care medicine, both of which commonly use electronic patient data management systems. Being able to use machine learning on this data requires that it is not simply stored but readily found, accessible, interoperable and reusable in accordance with the FAIR-principles. Analysis utilises a variety of supervised and unsupervised learning methods leading, amongst other things, to hypotheses for prospective randomised trials, analysis of rare complications, risk stratification and development of decision support tools. The aim of this article is to provide an overview of methods and application of machine learning in anaesthesiology and intensive care medicine. In addition, potential pitfalls associated with the technology and possible solutions are discussed. In anaesthesiology and intensive care medicine – as in other areas of medicine – machine learning can help provide individualised care with the aim of avoiding complications and increasing the quality of care provided.

### Einleitung

„Ich selbst habe eine App auf dem Handy, die mit 20 oder 30 Fragen Diagnosen genauer trifft als viele Ärzte, weil sie auf so viele Studien und Informatio-

nen zurückgreifen kann, wie es kein Arzt alleine kann.“ Diese Anmerkung des Bundesgesundheitsministers Jens Spahn zu einer auf künstlicher Intelligenz (KI) gestützten Gesundheitsplattform zeigt, dass KI mittlerweile Einzug in Gesundheitspolitik und Gesellschaft gefunden hat. Viel verwendete Schlagworte wie „künstliche Intelligenz“, „maschinelles Lernen“ oder „Big Data“ sind häufig nicht klar definiert und schüren Ängste. Diese betreffen neben Datenschutz und Schutz der Privatsphäre vor allem auch die Frage, ob der Mensch in Zukunft überflüssig sein wird. Können KI-Systeme Anästhesisten ersetzen? Schließlich werden in der Anästhesiologie durch den Einsatz elektronischer Patientendatenmanagementsysteme enorme Datenmengen generiert, die prädestiniert sind für die Anwendung von KI-Verfahren. Der Einsatz dieser neuen Methoden ermöglicht eine strukturierte Datenerfassung und -analyse sowie deren systematische Nutzung. KI-Verfahren versprechen, Komplikationen zu vermeiden, Abläufe effizient zu gestalten und Ressourcen optimal zu nutzen. Damit besitzt der verantwortungsvolle Einsatz von KI in Anästhesiologie und perioperativer Medizin das Potenzial, die Patientenversorgung zu verbessern. Ziel dieses Artikels ist es, dem Kliniker eine Einführung in die neuen Technologien zu geben, um anschließend auf spezielle Anwendungsmöglichkeiten in Anästhesiologie und Intensivmedizin einzugehen.

### Einführung: Künstliche Intelligenz in Anästhesiologie und Intensivmedizin

Mit Hilfe künstlicher Intelligenz (KI) sollen Entscheidungsstrukturen des menschlichen Denkens imitiert werden. Dazu entwickelt man Software, die zu eigendynamischem Verhalten fähig ist, und bedient sich dafür Techniken des maschinellen Lernens.

Durch den Einsatz von KI soll basierend auf großen Datenmengen eine personalisierte Behandlung für das Individuum gefunden werden. Diese großen Datensätze müssen bestimmte Voraussetzun-

gen erfüllen, damit sie mit Methoden des maschinellen Lernens auswertbar sind. Deswegen gehen wir im Folgenden zunächst auf die Datenstruktur ein, um dann Techniken des maschinellen Lernens zu erläutern.

### Datenstruktur

#### Big Data

Der nicht strikt definierte Begriff „Big Data“ beschreibt große, teils unübersichtliche Datenmengen, die nicht mit traditionellen Datenbank-Formaten zu verwalten oder mit herkömmlichen statistischen Methoden zu analysieren sind [2], und zusätzlich die Methoden zu deren Verarbeitung und Auswertung. „Big Data“ wird charakterisiert durch das sogenannte 5-V-Konzept: „Volume“ beschreibt die reine Quantität der Daten, „Velocity“ die Geschwindigkeit, mit der sie erhoben werden, „Variety“ die Diversität der Datenquellen, „Veracity“ die Richtigkeit oder Sinnhaftigkeit und „Value“ den Wert der Daten im ökonomischen Sinn [3,4].

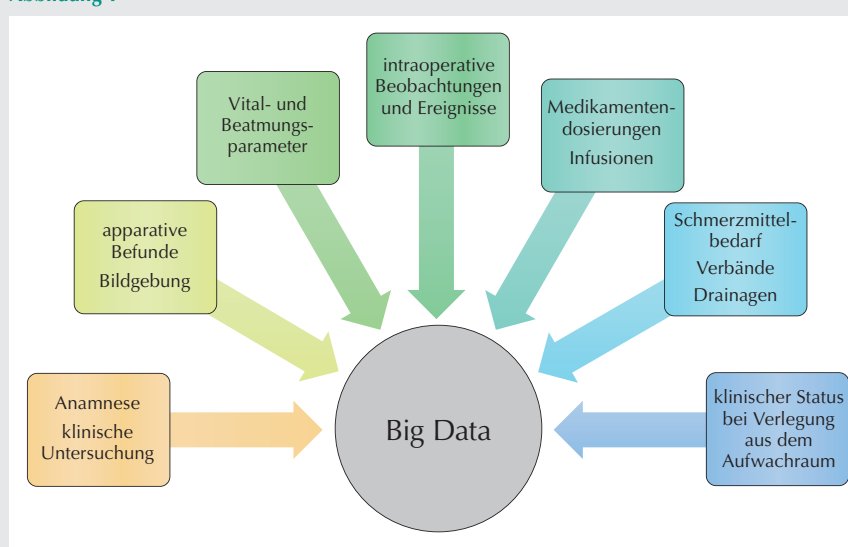
Bezogen auf die Medizin bedeutet dies: Die meisten Daten, die früher ausschließlich auf Papier erhoben wurden, sind heute in digitaler Form verfügbar. Die Erfassung von Vitalparametern und Laborwerten, klinischen Beobachtungen

sowie apparativen Befunden führt dazu, dass große Datenmengen („Volume“) mit teilweise hoher Geschwindigkeit, zum Beispiel bei kontinuierlich erhobenen Vitalparameterdaten („Velocity“), erhoben werden. Die Daten liegen in unterschiedlichsten Formaten – unter anderem auch als Text-, Audio- oder Videodateien („Variety“) – und in unterschiedlicher Auflösung, Genauigkeit und Korrektheit („Veracity“) vor. Wieviel Wert („Value“) nicht nur aus ökonomischer Sicht, sondern vor allem im Sinne einer besseren Behandlung des Patienten man letztendlich daraus schöpfen kann, hängt von der Qualität der Daten, deren Speicherung und Verarbeitung sowie dem zur Verfügung stehenden Fachwissen ab (Abb. 1).

#### Smart Data

Der Begriff „Smart Data“ fokussiert auf das fünfte V – „Value“ und beinhaltet die Integration fachspezifischen Wissens in den Datensatz, um aus Rohdaten verwertbare Informationen zu gewinnen [5]. Basis hierfür ist das vierte V – die „Veracity“. Diese wird zum einen gewährleistet durch hohe Präzision. Im klinischen Kontext entspricht die Präzision dem positiven prädiktiven Wert, also der Wahrscheinlichkeit, dass ein Patient mit

Abbildung 1



Von der Prämedikation über den intraoperativen Verlauf bis zur Verlegung aus dem Aufwachraum werden in der Anästhesiologie große Datenmengen generiert, die aufgrund ihrer Quantität und Diversität „Big Data“ sind.

positivem Testergebnis auch wirklich an der entsprechenden Erkrankung leidet. Die zweite Komponente, die Accuracy, übersetzt mit „Richtigkeit“, sagt aus, wie nahe der vorhergesagte und der richtige Wert beieinander liegen, ist also ein Maß für die korrekte Vorhersage eines Algorithmus. Hohe Präzision und hohe Accuracy steigern die Veracity der Daten. Nur aus korrekten Daten kann man Nutzen ziehen, wenn klinisches Wissen implementiert wird. Je nachdem, welcher Art dieser Nutzen ist, werden vier verschiedene Typen von Smart Data unterschieden: deskriptiv (beschreibend), prädiktiv (vorhersagend), präskriptiv (festlegend) und kognitiv (erkennend).

Zum Beispiel kann rein deskriptiv dargestellt werden, wie viele Patienten aus einem Kollektiv an PONV leiden. Der prädiktive Wert besteht dann in einer präoperativen Vorhersage des Risikos für PONV eines Patienten. Präskriptiv werden dann Vermeidungsstrategien für PONV entwickelt und kognitiv können die neuen Erkenntnisse dazu verwendet werden, Patienten und Ärzten die Ursachen, Auswirkungen und Prävention von PONV zu erklären [3].

### Patientendatenmanagementsysteme und medizinische Datenbanken

Die automatische Aufzeichnung von Narkosedaten wurde von einzelnen

Zentren bereits Ende der 1970er Jahre praktiziert und ist heute in vielen Operationssälen zur Routine geworden. In der Intensivmedizin hat sich das PDMS (Patient Data Management System) noch nicht flächendeckend durchgesetzt. Laut einer aktuellen Umfrage nutzen nur insgesamt 40% aller Intensivstationen ein digitales Dokumentationssystem, bei den Universitätskliniken beträgt der Anteil immerhin 58%. Dabei ist nicht nur die digitale Dokumentation, Archivierung und Auswertung von Vitaldaten, Verlaufseinträgen, Medikation und Interventionen essenziell. Das PDMS sollte möglichst mit allen Komponenten des Krankenhausinformationssystems (KIS) kommunizieren und auf Befunde aus anderen Informationssystemen, zum Beispiel Laborwerte und radiologische Befunde, zugreifen können [6]. Diese Daten aus unterschiedlichen Quellen werden dann in einem Data Warehouse, einer speziellen optimierten Datenbank, zusammengeführt, gespeichert, gefiltert, neu organisiert, strukturiert und verdichtet, um sie dann den entsprechenden Anwendungen zuführen zu können [7] (Abb. 2).

Für wissenschaftliche Zwecke wie der Entwicklung von Algorithmen beziehungsweise der Verbesserung und Korrektur vorhandener Systeme sind spezielle Datenbanken nötig. Ein seit

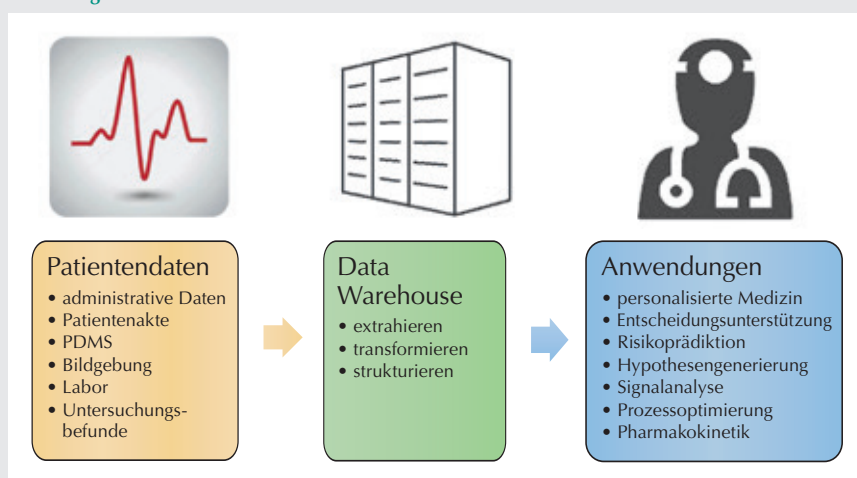
längerem erfolgreich betriebenes Projekt ist die MIMIC-Datenbank, eine für wissenschaftliche Zwecke frei zugängliche Intensivdatenbank aus Boston, aus der bereits unzählige Publikationen hervorgegangen sind. MIMIC enthält unter anderem Vitaldaten, Laborwerte, Diagnosen und Prozeduren von Intensivpatienten über mehr als zehn Jahre [8]. Auch in Deutschland wurde in jüngster Zeit eine Reihe von Konsortien im Rahmen der Medizininformatik-Initiative ([www.medizininformatik-initiative.de](http://www.medizininformatik-initiative.de)) gegründet, um Daten aus Krankenversorgung und Forschung zusammenzuführen. Hierzu zählen beispielsweise DIFUTURE (Data Integration for Future Medicine), HiGHmed (Heidelberg – Göttingen – Hannover Medical Informatics), MIRACUM (Medical Informatics for Research and Care in University Medicine) und SMITH (Smart Medical Information Technology for Healthcare). Weitere sind in Planung. Entscheidend für eine effektive gemeinsame Nutzung von Daten aus verschiedenen Zentren sind eine einheitliche Nomenklatur und einheitliche Datenformate. Hier ist vor allem SNOMED-CT zu nennen (systematized nomenclature of human and veterinary medicine clinical terms), das eine universelle Nomenklatur für Symptome, Befunde, Diagnosen und Prozeduren bildet [9]. Um die Daten aus verschiedenen Datenbanken in ein gemeinsames Format zu überführen, wird häufig das OMOP (observational medical outcomes partnership) Common Data Model verwendet [10]. Eine solche Vereinheitlichung ermöglicht einen standortübergreifenden Datenaustausch sowie eine systematische Verarbeitung von Daten aus verschiedenen Zentren. Durch die Zusammenführung von Daten aus mehreren Zentren erhält man auch für seltene Erkrankungen und Komplikationen umfangreiche Informationen, die die Versorgung der betroffenen Patienten deutlich verbessern kann.

### Künstliche Intelligenz zur Datenanalyse

#### Künstliche Intelligenz

Künstliche Intelligenz in ihrer ursprünglichen Definition beinhaltet die Nachbildung menschlichen Verhaltens

Abbildung 2



Strukturierte und unstrukturierte Daten müssen zusammengeführt und neu organisiert werden, um sie den entsprechenden Anwendungen zuführen zu können. Dies geschieht mit Hilfe spezieller, optimierter Datenbanken, den „Data Warehouses“.

mit Methoden der Mathematik und Informatik. KI-Systeme imitieren Entscheidungsstrukturen des Menschen [11]. Sie müssen lernfähig sein und mit einem gewissen Maß an Unsicherheit umgehen können. Maschinelles Lernen stellt einen Teilbereich der KI dar, die Begriffe „künstliche Intelligenz“ und „maschinelles Lernen“ werden jedoch häufig fälschlicherweise synonym verwendet.

Maschinelles Lernen

Ob Werbung oder Bildersuche im Netz, Spamfilterung von E-Mails oder das Entsperren des Smartphones über Gesichtserkennung – im Alltag wird maschinelles Lernen mittlerweile breit eingesetzt. Dabei handelt es sich um komplexe statistische Verfahren, die der Erkennung von Mustern, der Erstellung von Vorhersagen und der Hypothesengenerierung aus Datensätzen dienen, was mittels automatisierter Algorithmen erfolgt. Im Gegensatz zu herkömmlichen statistischen Methoden erstellt das künstliche System auf diese Weise Modelle, die nicht im Voraus programmiert worden sind, und es ist damit in der Lage, eine gestellte Aufgabe immer besser lösen zu können, sich also kontinuierlich zu optimieren [4], was als „Lernen“ bezeichnet wird.

Überwachtes Lernen

Meist werden überwachte maschinelle Lernverfahren vor allem in den Bereichen Risikoberechnung und Entscheidungsunterstützung angewendet. Beim überwachten Lernen (supervised learning) wird der angewendete Algorithmus mit einem Trainingsdatensatz hinsichtlich bekannter Ergebnisse trainiert, das heißt, Input und Output müssen bekannt sein. Dazu wird ein möglichst großer Trainingsdatensatz verwendet, der sogenannte Classifier, innerhalb dessen Regeln über das Verhältnis zwischen den einzelnen Datenelementen erkannt werden sollen. Ein weiterer Datensatz dient anschließend zur Validierung, und in einem dritten, dem Testdatensatz, wird die Vorhersagekraft des ermittelten Algorithmus bestätigt [11,12]. Dabei erfolgen Training und Validierung meist aus einem Datensatz, während der Testdatensatz ein davon unabhängiger, externer Datensatz ist. Diese „Kontrolle“

erfolgt jedoch nur während der Trainingsphase, im Produktivbetrieb wird das System nicht aktiv überwacht. Daher ist es wichtig, eine Überanpassung, ein sogenanntes „Overfitting“ zu vermeiden. Ein Zeichen dafür ist ein Ansteigen der Fehler in den Testdatensätzen, während die Anzahl der Fehler in den Trainingsdatensätzen fällt. Das heißt, das Modell liefert gute Prognosen in Bezug auf den Trainingsdatensatz, unbekannte Daten werden jedoch nicht mehr richtig zugeordnet, weil der Algorithmus zu genau auf den Trainingsdatensatz abgestimmt ist und etwaige Abweichungen zwischen den Datensätzen somit zu Fehlern führen [12]. Werden dagegen zu wenig Trainingsdaten zur Verfügung gestellt, kann es umgekehrt zu einem „Underfitting“ kommen. Beides gilt es durch ausreichendes Testen zu vermeiden. Methoden, die beim überwachten Lernen angewendet werden, sind zum Beispiel Entscheidungsbäume, neuronale Netze, Bayessche Netze oder Support Vector Machines.

Unüberwachtes Lernen

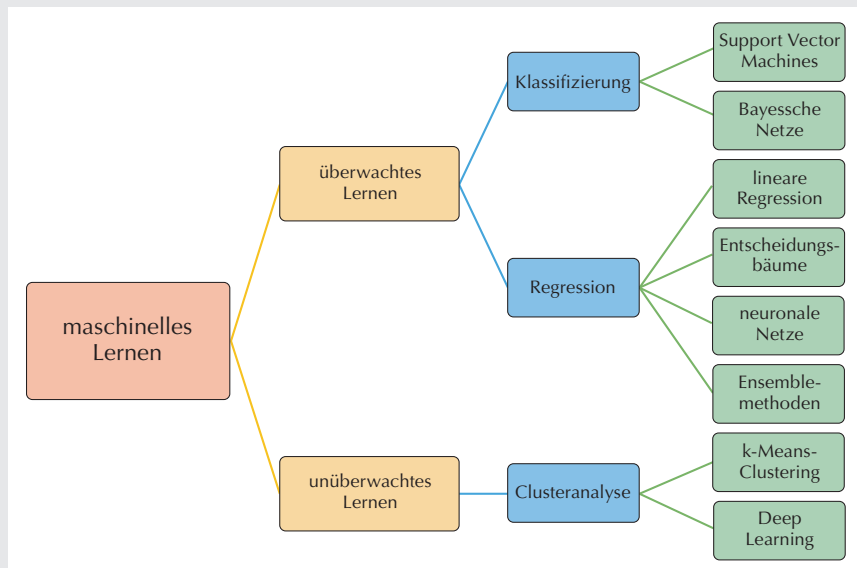
Im Gegensatz zum überwachten Lernen benutzt man beim unüberwachten Lernen (unsupervised learning) keine gekennzeichneten Daten, und der Output ist nicht bekannt [12]. Dadurch ist es nicht möglich, den Algorithmus zu trainieren. Stattdessen versuchen unüberwachte Verfahren, eine Struktur in Daten zu erkennen. Dazu bedient man sich meist Clustering-Methoden, die einzelne Beobachtungen bestimmten Teilmengen (Cluster) zuordnen, die in sich ähnliche und von anderen verschiedene Eigenschaften aufweisen. Erst die Interpretation etwaiger identifizierter Strukturen übernimmt dann der Mensch. Ein klinisches Beispiel hierfür ist die Einteilung von Herzinsuffizienz-Patienten mit erhaltener Pumpfunktion (HF-pEF), einer sehr heterogenen Patientengruppe. Die einzelnen Phänotypen konnten mittels Clustering in drei Gruppen eingeteilt werden, die sich in klinischen Charakteristika, Hämodynamik und Outcome unterscheiden. Dieses Gruppenmodell

Tabelle 1  
Häufig angewendete Methoden des maschinellen Lernens.

Methode	Beschreibung
Entscheidungsbaum (decision tree)	Dient der automatischen Klassifikation von Objekten. Besteht aus Wurzelknoten, inneren Knoten und Blättern. Jeder Knoten steht für eine logische Regel, jedes Blatt für eine Antwort.
Random Forest	Klassifikationsverfahren, das sich aus vielen verschiedenen Entscheidungsbäumen zusammensetzt.
bestärkendes Lernen (Reinforcement Learning)	Legt fest, wie ein zu eigendynamischem Verhalten fähiges Computerprogramm handeln muss, um Belohnungen zu maximieren.
Bayessche Netze	Erstellen graphische Modelle zufälliger Variablen und ihrer bedingten Abhängigkeiten.
Support Vector Machines	Dienen der Klassifikation und Regression. Anhand von Lernbeispielen, die in eine von zwei vorgegebenen Kategorien einzuordnen sind, erstellt ein Support-Vector-Machine-Algorithmus ein Modell, das vorhersagt, in welche der Kategorien ein neues Beispiel einzuordnen ist.
künstliche neuronale Netze	Bestehen aus miteinander verbundenen künstlichen Neuronen, die Informationen verarbeiten.
Deep Learning	Neuronale Netze mit zahlreichen verborgenen Schichten (deep layers), die dem neuronalen Netz eine gewisse „Tiefe“ verleihen.
Clustering	Zusammenfassung von Methoden des unüberwachten Lernens. Hier „weiß“ das System nicht, was es erkennen soll und teilt die Beobachtungen in verschiedene Kategorien (Cluster) ein.
Ensemblemethoden	Metaalgorithmen, die verschiedene Methoden des maschinellen Lernens kombinieren.



Abbildung 3



Maschinelles Lernen lässt sich in überwachtes und unüberwachtes Lernen einteilen. Domänen des überwachten Lernens sind Klassifizierung und Regression, während unüberwachtes Lernen in erster Linie Clusteranalysen beinhaltet.

kann zur Risikostratifizierung verwendet werden [13].

Einen Überblick über maschinelles Lernen und häufig angewendete Methoden geben Tabelle 1 und Abbildung 3.

### Anwendungsbereiche der Künstlichen Intelligenz in Anästhesiologie und Intensivmedizin

KI eröffnet zahlreiche Anwendungsmöglichkeiten auf dem Gebiet der personalisierten Medizin, der Entscheidungsunterstützung und Risikoprädiktion. Weitere Anwendungsgebiete sind die Signalanalyse und die Erstellung pharmakokinetischer Modelle. Im Bereich der epidemiologischen Forschung bietet KI ein großes Potenzial in der Untersuchung von seltenen Erkrankungen und Komplikationen sowie bei der Hypothesengenerierung für prospektiv randomisierte klinische Studien.

### Auswertung klinischer Studien, Hypothesengenerierung und Erfassung seltener Komplikationen

Klassische retrospektive Studien mit manueller Sichtung der Akten einer begrenzten Anzahl von Patienten sind

arbeitsaufwändig und unterliegen einer Reihe von Limitationen wie zum Beispiel einer unvollständigen Erfassung von Störfaktoren, widersprüchlichen Angaben in den Patientenakten und einer unterschiedlichen Bewertung klinischer Ereignisse [14]. Mittlerweile werden retrospektive Analysen anhand großer Datenmengen mit in der Regel mehreren tausend Patienten durchgeführt. Dabei kommen oft automatisierte Analyseverfahren zum Einsatz, die dem maschinellen Lernen zuzuordnen sind. Dadurch wurden retrospektive Kohorten- und Fall-Kontroll-Studien weitestgehend verdrängt [15]. Mittels Analysen großer Datenbanken generiert man außerdem Hypothesen, die dann in prospektiven randomisierten klinischen Studien weiter untersucht werden. Dies ermöglicht es, genaue Einschlusskriterien zu definieren [14]. Bailly et al. untersuchten zum Beispiel die Wirksamkeit einer antimykotischen Therapie bei Sepsis-Patienten mit Hilfe einer Intensiv-Datenbank [16]. In einer gezielten randomisierten Doppelblindstudie wurde die Hypothese dann weiter überprüft [17].

In prospektiven Untersuchungen mit hohen Fallzahlen erleichtern Methoden des

maschinellen Lernens die Auswertung und sind konventionellen statistischen Methoden überlegen, gerade wenn eine Vielzahl an Subgruppen oder der Einfluss multipler Faktoren analysiert werden soll. Beispiel hierfür ist eine Untersuchung an über 22.000 Patienten, in der gezeigt wurde, dass der Gebrauch von Muskelrelaxanzien die Rate an postoperativen pulmonalen Komplikationen erhöht [18].

Seltene Erkrankungen oder Ereignisse, für die in randomisierten klinischen Studien extrem hohe Fallzahlen erforderlich wären, die nahezu unmöglich zu erreichen sind, können mit Hilfe großer Datensätze aus möglichst vielen Kliniken analysiert werden. Vor allem in der Anästhesie betrifft dies Patientengruppen wie Kinder oder Schwangere, bei denen es nur wenige Ergebnisse aus randomisierten kontrollierten Studien gibt. Beispiele hierfür sind das Risiko für neuraxiale Hämatome in der geburts-hilflichen Anästhesie bei Thrombozytopenie [19] oder die Identifikation von Prädiktoren für das Versagen einer Larynxmaske bei Kindern, die sich deutlich von denen Erwachsener unterscheiden. Um 102 Kinder mit Larynxmaske zu identifizieren, die sekundär intubiert werden mussten, waren Daten von 11.910 Kindernarkosen nötig [20].

Nicht nur die Formulierung wissenschaftlicher Fragestellungen, die Entwicklung geeigneter Studiendesigns oder die Auswertung großer Datenmengen werden mit Hilfe von Methoden des maschinellen Lernens ermöglicht. Auch bei der Literaturrecherche helfen automatisierte Verfahren, die Suche zu ökonomisieren, was aufgrund der zunehmenden Informationsflut im Bereich der medizinischen Forschung notwendig ist. Data-Mining, die Anwendung automatisierter Verfahren zur Extraktion verwertbarer Informationen aus unübersichtlichen Datenmengen, ist auch für die Erstellung klinischer Leitlinien vonnöten. Die aktuellen Empfehlungen der ESA zur präoperativen Evaluation nicht-kardiochirurgischer Patienten wurden unter Verwendung von Data-Mining-Werkzeugen wie „PubReMiner“ oder „TerMine“ erstellt [21].

## Risiko-Prädiktion und Entscheidungsunterstützungssysteme

Die genaue Kenntnis des individuellen Risikos bezüglich des Behandlungserfolges ist essenziell, um eine für den jeweiligen Patienten maßgeschneiderte und damit personalisierte Anästhesie durchzuführen. Ziel ist, die perioperative Morbidität und Mortalität zu minimieren, was allgemein als „Kernkompetenz des Anästhesisten“ erachtet wird [22].

Risikokalkulatoren finden daher im klinischen Alltag zunehmend Verwendung. Es handelt sich dabei um komplexe Algorithmen, denen entweder stochastische Modelle oder Methoden des maschinellen Lernens zugrunde liegen [23] und die auf der Basis großer Datenbanken entwickelt werden. Sie finden mittlerweile in vielen Bereichen der Medizin Anwendung und haben auch Einzug in Leitlinienempfehlungen gehalten. Die ESA empfiehlt beispielsweise in ihren aktuellen Leitlinien zur präoperativen Evaluation den NSQIP (The National Surgery Quality Improvement Project)-Risikokalkulator zur Abschätzung des kardialen Risikos [21]. In die Entwicklung dieses Tools gingen klinische Daten aus 393 Krankenhäusern und von insgesamt über 1,4 Millionen Patienten ein. Entwickelt wurde der NSQIP-Risikokalkulator als Entscheidungshilfe für Patienten und Chirurgen. Basierend auf den präoperativ bekannten Risikofaktoren wurden Regressionsmodelle für postoperatives Outcome entwickelt. Dazu zählen kardiale, pulmonale oder infektiöse Komplikationen sowie das Auftreten einer Thrombose oder das Risiko, sich einer Re-OP unterziehen zu müssen. Das Risiko des einzelnen Patienten, eine oder mehrere dieser Komplikationen zu erleiden, wird mit Hilfe des NSQIP als webbasiertes Werkzeug berechnet und kann bei der Entscheidung für oder gegen eine OP mit einbezogen werden [24].

### Beispiele aus der Anästhesie

Bedient sich der NSQIP noch klassischer Regressionsmodelle, so werden in der aktuellen Literatur im Bereich

der Risikostratifizierung zunehmend Algorithmen des maschinellen Lernens verwendet. Im perioperativen Bereich können so zum Beispiel das Delirrisiko, das PONV-Risiko oder das Risiko einer Verschlechterung im Aufwachraum ermittelt werden [25–27]. Dabei verwendet zum Beispiel die Gruppe um Olsen einen Random Forest Classifier, um frühzeitig eine klinische Verschlechterung von Patienten im Aufwachraum zu erfassen [27]. In der Arbeit von Davoudi zur Ermittlung des Delirrisikos wurden verschiedene Verfahren verglichen, unter anderem Random-Forest-Modelle, Support Vector Machines und neuronale Netze [25]. Die Übergänge von Modellen zur Risikoberechnung und zur Entscheidungsunterstützung sind dabei fließend. Die im PDMS aufgezeichneten und analysierten Daten werden zunächst benutzt, um Risiken zu berechnen. Die Risikoberechnung ist dann von klinischer Relevanz, wenn in einem weiteren Schritt über im System hinterlegte Algorithmen dem Arzt Handlungsempfehlungen gegeben werden [28]. So stellten Kappen und Mitarbeiter bei der Implementierung ihres On-Screen-Tools zur Bestimmung des PONV-Risikos fest, dass die Anzeige des Risikos allein nicht zu einer geringeren PONV-Inzidenz führte. Erst der direkte Ansatz mit konkreten Therapieempfehlungen konnte das Behandlungsergebnis verbessern [26]. Um eine unmittelbare klinische Konsequenz aus der Risikoberechnung ableiten zu können, sind Echtzeitverfahren notwendig, beispielsweise für die Steuerung der Volumentherapie bei großen abdominal-chirurgischen Eingriffen [29] oder für die Vorhersage einer intraoperativen Hypotension aus der Form der arteriellen Druckkurve [30].

### Beispiele aus der Intensivmedizin

Computerbasierte Entscheidungsunterstützungssysteme, beispielsweise zur Festlegung der optimalen Antibiotikatherapie bei Intensivpatienten, haben in den letzten Jahren Einzug in den klinischen Alltag gehalten [31]. Auch Algorithmen zur Risikoprädiktion werden zunehmend entwickelt und angewendet. Speziell für die Intensivmedizin wurde von Piracchio und Mitarbeitern rein

auf Basis des maschinellen Lernens ein „Super ICU Learner“-Algorithmus erstellt, der die Krankenhausmortalität vorhersagt. Dieser Algorithmus war in seiner Prognose zuverlässiger als herkömmliche Scores [32].

In der pädiatrischen Intensivmedizin wurde ein Studienprotokoll vorgestellt, in dem kardiorespiratorische Signale von 250 Frühgeborenen erfasst und mit Hilfe eines Maschinenlern-Algorithmus analysiert werden sollen. Anhand dieser Daten soll der geeignete Extubationszeitpunkt bei intubierten Frühchen vorhergesagt werden [33]. Ein vielversprechendes Reinforcement-Learning-Modell wurde zur Therapiesteuerung der Sepsis entwickelt. In der Analyse eines retrospektiven Datensatzes hatte sich dabei gezeigt, dass die Überlebenswahrscheinlichkeit der Patienten höher war, wenn die vom Kliniker durchgeführte Therapie mit der vom System vorgeschlagenen übereinstimmte [34].

Diese Beispiele stehen exemplarisch für eine ganze Reihe von Entscheidungsunterstützungssystemen, die im Rahmen von Studien erprobt werden. Die routinemäßige Implementierung dieser Algorithmen in PDMS-Systeme bietet noch Potenzial für die Zukunft.

### Weitere Anwendungsgebiete

Am weitesten fortgeschritten ist man zurzeit auf dem Gebiet der Signalanalyse, vor allem bei der EKG-Auswertung. Rhythmusstörungen können mittlerweile mittels Smartwatches vom Patienten zu Hause detektiert werden [35]. Auch Narkosetiefemonitoring und die intraoperative Analyse somatosensibel evozierter Potenziale erfolgten mit Hilfe von maschinellen Lernalgorithmen [36,37].

Maschinenlern-Algorithmen werden auch eingesetzt bei der Entwicklung von Closed-Loop-Systemen für die Anästhesie, die Hypnose, Analgesie und Muskelrelaxation automatisch steuern. Aufsehen in der Öffentlichkeit erregte der kombinierte Einsatz des automatisierten Anästhesiesystems „McSleepy“ und des Da-Vinci-Operationsroboters [38]. Ein System, das zur Sedierung von Patienten für endoskopische Eingriffe entwickelt wurde (Sedasys), wurde je-

**Tabelle 2**

Anwendungsgebiete maschinellen Lernens in der Anästhesiologie und Intensivmedizin und Beispiele aus der Klinik.

Anwendungsgebiet	Beispiele
<b>Hypothesengenerierung</b>	Wirksamkeit der antimykotischen Therapie bei Sepsis [16]
<b>Erfassung seltener Komplikationen</b>	neuraxiale Hämatomate in der Geburtshilfe [19] Atemwegssicherung bei Kindern [20]
<b>Literaturrecherche</b>	ESA-Leitlinien zur präoperativen Evaluation nicht-kardiochirurgischer Patienten [21]
<b>Entscheidungsunterstützung</b>	Prädiktion von PONV [26] Therapiesteuerung der Sepsis [34]
<b>Risikoprädiktion</b>	NSQIP-Risikokalkulator [24] Mortalität auf Intensivstation (Super ICU Learner) [32]
<b>Signalanalyse</b>	Narkosetiefenmonitoring [36] Interpretation somatosensibler Potenziale [37]
<b>OMICS</b>	Biomarker für Sepsis [41] Aufwachzeiten nach Allgemeinanästhesie [42]
<b>Prozessoptimierung</b>	Optimierung der OP- und Aufwachraumkapazitäten [43] Transfusionslisten [44]
<b>pharmakokinetische Modelle</b>	Arzneimittelnebenwirkungen [46]

doch nach kurzer Zeit wieder vom Markt genommen, wohl in erster Linie aufgrund fehlender Akzeptanz bei Patienten, Ärzten und Pflegekräften [39]. Trotzdem steht die Entwicklung von Closed-Loop-Systemen nicht still. Studien an kleineren Patientenkohorten zeigen die technische Machbarkeit und liefern vielversprechende Ergebnisse [40]. Leider fehlen – vermutlich aufgrund des hohen technischen Aufwands und der Kosten – noch Untersuchungen an großen Patientenkollektiven, vor allem zu Patienten mit komplexen Vorerkrankungen.

Omics-Daten werden in der Intensivmedizin genutzt, um neue Biomarker für Sepsis zu finden, mit Hilfe derer man infektiöse und nicht-infektiöse Inflammation unterscheiden sowie das klinische Behandlungsergebnis und das Ansprechen auf bestimmte Therapiemaßnahmen vorhersagen kann [41].

In einer Arbeit aus der Anästhesie konnte gezeigt werden, dass die postoperativen Aufwachzeiten der Patienten abhängig von ihrem Genotyp sind, sodass man Patienten mit verlängerter Aufwachzeit nach Allgemeinanästhesie bereits im Voraus identifizieren könne [42].

Auch die Optimierung klinikinterner Abläufe und logistischer Prozesse erfolgt

mit Hilfe von maschinellen Lernalgorithmen, beispielsweise bei der Abstimmung von OP- und Aufwachraumkapazitäten, deren optimale Nutzung wesentlich die Kosten reduzieren und die Personalplanung erleichtern [43]. Weitere Einsatzgebiete sind die Errechnung des Bedarfs an Blutkonserven und die Erstellung von Transfusionslisten für bestimmte Eingriffe [44].

KI und maschinelles Lernen haben auch Einzug in die pharmakologische Forschung gehalten. Mittels eines Deep-Learning-Modells kann bei einer „target-controlled infusion“ (TCI) mit Propofol und Remifentanyl der Bispektral-Index (BIS) vorhergesagt werden [45]. Auch Modelle zur Vorhersage von Arzneimittelnebenwirkungen wurden bereits entwickelt, die zu einer erhöhten Patientensicherheit beitragen können [46].

Eine Zusammenfassung der Anwendungsgebiete zeigt Tabelle 2.

### Personalisierte Medizin

Die Medizin des 21. Jahrhunderts vollzieht einen Wandel von der rein „reaktiven“ Medizin hin zu einem prädiktiven, präventiven und personalisierten Ansatz. Dabei gilt es, Risikopatienten zu identifizieren, bevor Komplikationen

auftreten [47]. Dies ist vor allem vor dem Hintergrund relevant, dass weltweit jährlich 4,2 Millionen Menschen innerhalb von 30 Tagen nach einer Operation versterben. Somit liegt die postoperative Mortalität auf Platz 3 der häufigsten Todesursachen [48]. Im Bereich der Intensivmedizin wurden bereits Algorithmen entwickelt, die dazu dienen, Risikopatienten frühzeitig zu identifizieren. Masino und Mitarbeiter beispielsweise erstellten ein Modell, das Sepsis bei Neugeborenen vorhersagen kann, bevor sich klinische Zeichen entwickeln [49]. Für Erwachsene existiert ein Modell zur Vorhersage der Mortalität im kardiogenen und septischen Schock [50]. Mithilfe derartiger Werkzeuge gelingt es, lebensrettende Entscheidungen für den einzelnen Patienten frühzeitig zu treffen [50].

Die Etablierung zielgerichteter Präventionsmaßnahmen und die Entwicklung von individuellen, maßgeschneiderten Behandlungen versprechen eine spürbare Verbesserung der Versorgung vor allem von Patienten mit komplexen Krankheitsbildern [2]. Noch fehlen groß angelegte Multicenterstudien, die eine Verbesserung des Patientenoutcomes beim Einsatz von Methoden des maschinellen Lernens im Vergleich zu konventioneller Behandlung belegen. Große Datenmengen aus PDMS und elektronischen Patientenakten, die mit Hilfe von maschinellen Lernalgorithmen ausgewertet werden, können uns allerdings dabei helfen, die Mortalität zu senken, die Qualität der Behandlung zu verbessern und die Patientenzufriedenheit zu steigern [51,52].

### Probleme und Lösungsansätze

#### Risiken

Cabitza et al. fassen in ihrer Stellungnahme aus dem Jahr 2017 die größten Befürchtungen der Kritiker maschinellen Lernens zusammen. Überspitzt formuliert werden die Risiken maschinellen Lernens folgendermaßen charakterisiert: Durch die Anwendung computergestützter Algorithmen verlieren Ärzte die Fähigkeit, eigenständig Diagnosen und Therapiepläne zu erstellen. Die verwen-



deten Algorithmen sind aus verschiedenen Gründen fehleranfällig. Diese Fehler können aber nicht erkannt werden, weil Maschinenlern-Algorithmen als „black boxes“ undurchschaubar und damit einer Fehleranalyse nicht zugänglich sind [53].

Das sogenannte „De-Skilling“, das dazu führt, dass Ärzte dazu neigen, fehlerhafte Entscheidungen des Systems einfach zu übernehmen, obwohl sie mit ihrer klinischen Erfahrung eigentlich in der Lage wären, die richtige Entscheidung zu treffen, ist als Problem nicht von der Hand zu weisen. So konnte die Anwendung eines computerassistierten Systems bei unerfahrenen Ärzten zu einer Erhöhung der Anzahl korrekter Diagnosen führen. Erfahrenere Diagnostiker jedoch schlossen sich im Zweifel eher einer vom System vorgeschlagenen falschen Diagnose an und hätten ohne Nutzung des Systems eine höhere Genauigkeit in ihrer Diagnosestellung erzielt [53,54].

Die Fehleranfälligkeit von Algorithmen rührt daher, dass sie im Gegensatz zum Menschen den Kontext in bestimmten Situationen nicht erkennen oder falsch einordnen. Dies geschieht vor allem dann, wenn ein Trainingsdatensatz von den realen Bedingungen abweicht. So sollte ein KI-System zum Beispiel Pferde auf Bildern erkennen. Etwa ein Fünftel der Bilder des Trainingsdatensatzes waren mit einer Quellenangabe versehen. Entfernte man den Tag mit der Quellenangabe, so wurde das Pferd nicht mehr als solches erkannt [55]. Hinzu kommt eine „intrinsische Unsicherheit“ in der Medizin [53]. Gehen Daten in einen Algorithmus ein, die bereits zuvor von Menschen bewertet und beurteilt wurden, so wird dieser von subjektiven Faktoren beeinflusst, was auch eine Fehlerursache sein kann.

Die mangelnde Erklärbarkeit, „explainability“ genannt, führt dazu, dass die von Maschinenlern-Algorithmen ausgegebenen Ergebnisse oder Empfehlungen nicht nachvollziehbar sind, sie sind undurchsichtig, sogenannte „black boxes“ [56,57]. Eine Fehleranalyse ist damit nahezu unmöglich. Hier wird man sich allerdings gerade als Anästhesist der

Tatsache bewusst, dass die moderne Anästhesie in ihrer Entwicklung deutlich gehemmt worden wäre, hätte man sämtliche „black boxes“ von Anfang an aus den Operationssälen verbannt [56]. Schließlich wurden Anästhetika verwendet, lange bevor man eine Vorstellung von ihren Wirkmechanismen hatte. Natürlich darf man heute nicht die ethischen und qualitativen Maßstäbe des 19. Jahrhunderts ansetzen. Daher erfordert der verantwortungsvolle Umgang mit maschinellem Lernen und künstlicher Intelligenz gewisse Voraussetzungen, vor allem was die Schulung des medizinischen Personals und die Art und Qualität der verwendeten Daten betrifft.

### Schulung

Elektronische Patientenakten und telemedizinische Konsultationen gehören zum Alltag in der klinischen Praxis. Obwohl die meisten Kliniker tagtäglich Informationen aus großen Datenbanken für die Routine nutzen, sind sie nur wenig vertraut mit den analytischen Methoden zu deren Auswertung. Mangelndes Verständnis der zugrundeliegenden Zusammenhänge sowie unkritische Verwendung der vorhandenen Daten können dabei zu Fehlinterpretationen und damit zu Fehldiagnosen und falschen Behandlungen führen. Wichtig ist daher eine Schulung der Ärzte und vor allem auch der klinischen Forscher im Umgang mit großen Datensätzen und elektronischen Dokumentations- und Entscheidungshilfen sowie eine gründliche Validierung der in Klinik und Wissenschaft eingesetzten Werkzeuge. Wissen und Verständnis für die klinischen Zusammenhänge sind auch im Krankenhaus der Zukunft unabdingbar.

### Diversität von Daten

Medizinische Daten zeichnen sich dadurch aus, dass sie heterogen sind, sowohl in strukturierter als auch in unstrukturierter Form vorliegen und aus unterschiedlichen Quellen stammen [7]. Strukturierte Daten aus elektronischen Patientenakten, zum Beispiel Vitalparameteraufzeichnungen aus dem PDMS, sind relativ leicht in Datenbanken zu

archivieren, zu strukturieren und zu analysieren. Traditionell enthalten Patientenakten jedoch noch viele unstrukturierte Daten, in erster Linie handschriftliche Aufzeichnungen von Ärzten und Pflegekräften oder Freitextbefunde von apparativen Untersuchungen. Gerade unleserliche handschriftliche Notizen beeinträchtigen die „veracity“ von Daten und führen zu Fehlinterpretationen. Man weiß seit längerem, dass ein einfacher Weg Verschreibungsfehler zu vermeiden darin besteht, digitale Eingaben statt handschriftlicher Anordnungen zu nutzen. Leider stoßen strukturierte Eingabemasken beim Klinikpersonal häufig auf Widerstand, weil es Freitextanordnungen gewohnt ist. Um strukturierte Daten zu erhalten ist es aber erforderlich, Freitexteingaben zu minimieren und vorgegebene Eingabefelder zu verwenden. Für die Akzeptanz bei Ärzten und Pflegepersonal ist es daher entscheidend, benutzerfreundliche Software und Eingabemasken zu entwickeln [2].

Bis sich strukturierte Eingabemasken flächendeckend durchgesetzt haben, muss immer noch Information aus unstrukturiertem Freitext extrahiert werden. Herkömmliche Methoden des Text-Minings beziehungsweise des Extrahierens von Informationen aus großen Freitextmengen, das auf statistischen und probabilistischen Modellen beruht, versagen etwa bei Schreibfehlern und mehrdeutigen Wörtern. Hinzu kommen grammatikalische Besonderheiten vor allem bei nicht-englischsprachigen Texten. Ein Beispiel dafür sind Verneinungen, die unterschiedliche Formen haben können („kann ausgeschlossen werden“, „frei von Beschwerden“, „Hinweise auf... gibt es nicht“) und für deren Erkennung neue Werkzeuge entwickelt oder vorhandene adaptiert werden müssen [58]. Natural Language Processing ist eine Domäne des Deep Learning. Damit gelingt es zunehmend, Bedeutungen und Zusammenhänge auch aus komplexen Texten zu extrahieren.

### Datenerhebung

Bei speziellen Messverfahren fehlen oft die Schnittstellen zum PDMS, zum Beispiel bei der Anwendung separater Systeme für multimodales Neuromoni-



toring oder erweitertes hämodynamisches Monitoring. Sind verschiedene Kliniken beteiligt, ist es wichtig, vor Studienbeginn die nötigen Schnittstellen zu identifizieren, eine Kompatibilität der verwendeten Systeme sicherzustellen und sich auf eine einheitliche Nomenklatur, eine einheitliche Definition klinisch relevanter Ereignisse und ein einheitliches Archivierungsformat zu einigen. Der technische und personelle Aufwand wird dabei manchmal unterschätzt. Die Harmonisierung der erhobenen Daten ist jedoch entscheidend für deren Verwertbarkeit [59]. Fachgesellschaften geben zunehmend Empfehlungen zu Datenerhebung und -management im Rahmen großer Multicenterstudien heraus [60]. Generell gilt für große Datensätze das FAIR-Prinzip: Die Daten sollen wiederzufinden (Findable), leicht zugänglich (Accessible), von verschiedenen Systemen zu bearbeiten (Interoperable) und wiederholt zu benutzen (Reusable) sein [61].

### Datenqualität

Sind Datensätze nicht komplett oder sind invalide Daten enthalten, so erhöht dies bei Anwendung konventioneller statistischer Methoden die Wahrscheinlichkeit eines Fehlers 2. Art. Big Data ist dagegen nicht so anfällig gegenüber inkompletten oder unpräzisen Daten [7].

Daten können auch mit unterschiedlichen Auflösungen gespeichert sein. So können zum Beispiel invasive Blutdrücke minütlich, alle 5 Sekunden oder etwa in einer Auflösung von 100 Hz gespeichert werden. Hier muss man einen Kompromiss zwischen Informationsverlust und hohem Speicherplatzbedarf finden.

Außerdem wird die Datenqualität häufig durch Artefakte beeinträchtigt. In einer Studie bei Kindernarkosen betrug die Artefaktrate bei nicht-invasiven Blutdrücken 5%, bei invasiven Blutdrücken sogar 7,3% [62]. Wichtig ist daher, abweichende Werte zu definieren und aus diesen abweichenden Werten die Artefakte zu identifizieren. Um dies zu erreichen, können aus dem Gebiet der KI-Technologien zum Beispiel rekurrente neuronale Netze (neuronale Netze, die Verbindungen innerhalb derselben oder auch zurückliegender Schichten besitzen und so über Rückkopplungsmechanismen verfügen) und Long-Short-Term-Memory-Techniken (neuronale Netze mit einer Art Erinnerung an frühere Erfahrungen) angewendet werden.

### Datenschutz und Datensicherheit

Um den aktuellen Datenschutzbestimmungen zu genügen und einen Missbrauch von Patientendaten zu verhindern, müssen Gesundheitsdaten anonymisiert werden [7]. Dies ist jedoch nicht immer ohne Qualitätsverlust möglich. Bestimmte Datensätze, z.B. Genomdaten, können auch gar nicht im eigentlichen Sinn anonymisiert werden. Um die Anonymität der Daten zu gewährleisten ohne sie für die medizinische Diagnostik unbrauchbar zu machen, müssen die Daten einerseits bereinigt werden, das heißt, unwichtige Informationen werden weggelassen. Andererseits kann ein

„Rauschen“, also nicht zum eigentlichen Datensatz gehörige Informationen hinzugefügt werden, um die Identifizierung zu erschweren (software-cluster.org, Beitrag vom 02.08.2017).

### Ethische Fragestellungen

Die Vielzahl an ethischen Fragestellungen zu diskutieren, die sich mit zunehmendem Einsatz von KI in der Medizin ergeben, würde den Rahmen dieses Artikels sprengen. So ist unter anderem die Haftungsproblematik im Falle der Schädigung eines Patienten nicht ausreichend geklärt. Darüber hinaus stellt sich aber auch die Frage, ob man dem Patienten KI-Systeme zur Diagnostik und Therapieplanung vorenthalten darf, wenn diese doch treffsicherer als der Arzt sind. Der Bundesverband digitale Wirtschaft behandelt derartige Fragestellungen in seinem Diskussionspapier „Mensch, Moral, Maschine. Digitale Ethik und künstliche Intelligenz“ (www.bvdw.org, abgerufen am 13.06.2019).

Ein weiterer Punkt ist die bereits oben erwähnte Tatsache, dass Maschinenlern-Algorithmen häufig sehr komplex und damit nicht erklärbar sind. Allerdings werden zunehmend Methoden entwickelt, die Input und Output analysieren und die Faktoren identifizieren können, die zur jeweiligen Aussage oder Empfehlung führen. Ein Beispiel für eine derartige Methode ist das sogenannte LIME-Modell (local interpretable model-agnostic explanations). Ein Trend geht auch dahin, zunehmend erklärbare Modelle zu verwenden [52,57].

### Ausblick

Die Medizin befindet sich im Wandel. Die Verfügbarkeit großer Datensätze und das Vorhandensein technischer Möglichkeiten für deren Speicherung und Auswertung eröffnen im Bereich der Anästhesiologie und Intensivmedizin neue Möglichkeiten. Epidemiologische Studien, Analyse des Behandlungserfolges und Risikostratifizierung werden durch Big Data-Methoden vereinfacht. Subgruppenanalysen ermöglichen es, gezielte prospektive Studien bei bestimmten Patientenkollektiven zu planen und durchzuführen. Durch strukturierte Aufzeichnung, Archivierung und Analyse von

Daten sowie gezieltes Zusammenführen von Datensätzen können mehr Daten zu Patienten mit seltenen Erkrankungen oder seltenen perioperativen Komplikationen erfasst werden. Ziel ist eine individuelle Risikoeinschätzung und optimierte Therapie für den einzelnen Patienten, eine Ökonomisierung der perioperativen Abläufe und eine Verringerung der Komplikationsraten.

### Literatur

1. Jungk R: Die Zukunft hat schon begonnen. Entmenslichung – Gefahr unserer Zivilisation. Bern/Stuttgart: Goldmann 1952. Amerikanische Ausgabe: Tomorrow is already here. New York: Simon and Schuster 1954
2. Raghupathi W, Raghupathi V: Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst* 2014;2:3
3. Hofer IS, Halperin E, Cannesson M: Opening the Black Box: Understanding the Science Behind Big Data and Predictive Analytics. *Anesth Analg* 2018;127:1139–1143
4. Bailly S, Meyfroidt G, Timsit JF: What's new in ICU in 2050: big data and machine learning. *Intensive Care Med* 2018;44:1524–1527
5. Funkner AA, Yakovlev AN, Kovalchuk SV: Data-driven modeling of clinical pathways using electronic health records. *Pr* 2017;121:835–842
6. Baumgartel M, Riessen R, John S: Digitization in Intensive Care Medicine. *Dtsch Med Wochenschr* 2019;144:436–441
7. Austin C, Kusumoto F: The application of Big Data in medicine: current implications and future directions. *J Interv Card Electrophysiol* 2016;47:51–59
8. Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, Ghassemi M, et al: MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035
9. Chu L, Kannan V, Basit MA, Schaefflein DJ, Ortuzar AR, Gloriosio JF, et al: SNOMED CT Concept Hierarchies for Computable Clinical Phenotypes From Electronic Health Record Data: Comparison of Intensional Versus Extensional Value Sets. *JMIR Med Inform* 2019;7:e11487
10. Maier C, Lang L, Storf H, Vormstein P, Bieber R, Bernarding J, et al: Towards Implementation of OMOP in a German University Hospital Consortium. *Appl Clin Inform* 2018;9:54–61
11. Dilsizian SE, Siegel EL: Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Curr Cardiol Rep* 2014;16:441
12. Lee S, Mohr NM, Street WN, Nadkarni P: Machine Learning in Relation to Emergency Medicine Clinical and Operational Scenarios: An Overview. *West J Emerg Med* 2019;20:219–227
13. Shah SJ, Katz DH, Selvaraj S, Burke MA, Yancy CW, Gheorghiade M, et al: Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation* 2015;131:269–279
14. Sessler DI: Big Data-and its contributions to peri-operative medicine. *Anaesthesia* 2014;69:100–105
15. Bhatt DL, Mehta C: Adaptive Designs for Clinical Trials. *N Engl J Med* 2016;375:65–74
16. Bailly S, Bouadma L, Azoulay E, Orgeas MG, Adrie C, Souweine B, et al: Failure of empirical systemic antifungal therapy in mechanically ventilated critically ill patients. *Am J Respir Crit Care Med* 2015;191:1139–1146
17. Timsit JF, Azoulay E, Schwebel C, Charles PE, Cornet M, Souweine B, et al: Empirical Micafungin Treatment and Survival Without Invasive Fungal Infection in Adults With ICU-Acquired Sepsis, Candida Colonization, and Multiple Organ Failure: The EMPIRICUS Randomized Clinical Trial. *JAMA* 2016;316:1555–1564
18. Kirmeier E, Eriksson LI, Lewald H, Jonsson Fagerlund M, Hoeft A, Hollmann M, et al: Post-anaesthesia pulmonary complications after use of muscle relaxants (POPULAR): a multicentre, prospective observational study. *Lancet Respir Med* 2019;7:129–140
19. Klumpner TT, Bauer ME, Kheterpal S: Current applications of big data in obstetric anaesthesiology. *Curr Opin Anaesthesiol* 2017;30:300–305
20. Mathis MR, Haydar B, Taylor EL, Morris M, Malviya SV, Christensen RE, et al: Failure of the Laryngeal Mask Airway Unique and Classic in the pediatric surgical patient: a study of clinical predictors and outcomes. *Anesthesiology* 2013;119:1284–1295
21. De Hert S, Staender S, Fritsch G, Hinkelbein J, Afshari A, Bettelli G, et al: Pre-operative evaluation of adults undergoing elective noncardiac surgery: Updated guideline from the European Society of Anaesthesiology. *Eur J Anaesthesiol* 2018;35:407–465

22. Boehm O: Morbidität und Sterblichkeit in der Anästhesiologie. Patienten-sicherheit als Kernkompetenz des Anästhesisten. *Anästh Intensivmed* 2019;60:488–500
23. Mansmann U, Rieger A, Strahwald B, Crispin A: Risk calculators-methods, development, implementation, and validation. *Int J Colorectal Dis* 2016;31:1111–1116
24. Bilimoria KY, Liu Y, Paruch JL, Zhou L, Kmiecik TE, Ko CY, et al: Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. *J Am Coll Surg* 2013;217:833–842
25. Davoudi A, Ebadi A, Rashidi P, Ozrazgat-Baslanti T, Bihorac A, Bursian AC: Delirium Prediction using Machine Learning Models on Preoperative Electronic Health Records Data. *Proc IEEE Int Symp Bioinformatics Bioeng* 2017;2017:568–573
26. Kappen TH, Vergouwe Y, van Wolfswinkel L, Kalkman CJ, Moons KG, van Klei WA: Impact of adding therapeutic recommendations to risk assessments from a prediction model for postoperative nausea and vomiting. *Br J Anaesth* 2015;114:252–260
27. Olsen RM, Aasvang EK, Meyhoff CS, Dissing Sorensen HB: Towards an automated multimodal clinical decision support system at the post anesthesia care unit. *Comput Biol Med* 2018; 101:15–21
28. Simpao AF, Tan JM, Lingappan AM, Galvez JA, Morgan SE, Krall MA: A systematic review of near real-time and point-of-care clinical decision support in anesthesia information management systems. *J Clin Monit Comput* 2017;31:885–894
29. Joosten A, Alexander B, Duranteau J: Clinical decision support system clears the way for perioperative goal directed therapy protocol adherence improvement. *Minerva Anestesiol* 2019;85:691–692
30. Hatib F, Jian Z, Buddi S, Lee C, Settels J, Sibert K, et al: Machine-learning Algorithm to Predict Hypotension Based on High-fidelity Arterial Pressure Waveform Analysis. *Anesthesiology* 2018;129:663–674
31. Tafelski S, Sacher A, Nachtigall I, Powollik R, Tamarkin A, Angermair S et al: Unterstützung der Antibiotikatherapie auf Intensivstationen in Deutschland. Analyse der Anwenderfreundlichkeit und Zufriedenheitsanalyse mit einem computerassistenten Stewardship-Programm. *Anästh Intensivmed* 2016;57:174–181
32. Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, van der Laan MJ: Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. *Lancet Respir Med* 2015;3:42–52
33. Shalish W, Kanbar LJ, Rao S, Robles-Rubio CA, Kovacs L, Chawla S, et al: Prediction of Extubation readiness in extremely preterm infants by the automated analysis of cardiorespiratory behavior: study protocol. *BMC Pediatr* 2017;17:167
34. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA: The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med* 2018;24:1716–1720
35. Dorr M, Nothruft V, Brasier N, Bosshard E, Djurdjevic A, Gross S, et al: The WATCH AF Trial: SmartWATCHes for Detection of Atrial Fibrillation. *JACC Clin Electrophysiol* 2019;5:199–208
36. Sadrawi M, Fan SZ, Abbod MF, Jen KK, Shieh JS: Computational Depth of Anesthesia via Multiple Vital Signs Based on Artificial Neural Networks. *Biomed Res Int* 2015;2015:536863
37. Fan B, Li HX, Hu Y: An Intelligent Decision System for Intraoperative Somatosensory Evoked Potential Monitoring. *IEEE Trans Neural Syst Rehabil Eng* 2016;24:300–307
38. Wehbe M, Arbeid E, Cyr S, Mathieu PA, Taddei R, Morse J, et al: A technical description of a novel pharmacological anesthesia robot. *J Clin Monit Comput* 2014;28:27–34
39. Goudra B, Singh PM: Failure of Sedasys: Destiny or Poor Design? *Anesth Analg* 2017;124:686–688
40. Zaouter C, Hemmerling TM, Mion S, Leroux L, Remy A, Ouattara A: Feasibility of Automated Propofol Sedation for Transcatheter Aortic Valve Implantation: A Pilot Study. *Anesth Analg* 2017;125: 1505–1512
41. Prucha M, Zazula R, Russwurm S: Sepsis Diagnostics in the Era of „Omics“ Technologies. *Prague Med Rep* 2018; 119:9–29
42. Xie S, Ma W, Shen M, Guo Q, Wang E, Huang C, et al: Clinical and pharmacogenetics associated with recovery time from general anesthesia. *Pharmacogenomics* 2018;19:1111–1123
43. Fairley M, Scheinker D, Brandeau ML: Improving the efficiency of the operating room environment with an optimization and machine learning model. *Health Care Manag Sci* 2019;22:756–767
44. Pendry K: The use of big data in transfusion medicine. *Transfus Med* 2015; 25:129–137
45. Lee HC, Ryu HG, Chung EJ, Jung CW: Prediction of Bispectral Index during Target-controlled Infusion of Propofol and Remifentanyl: A Deep Learning Approach. *Anesthesiology* 2018;128: 492–501
46. Dey S, Luo H, Fokoue A, Hu J, Zhang P: Predicting adverse drug reactions through interpretable deep learning framework. *BMC Bioinformatics* 2018;19:476
47. Van Poucke S, Zhang Z, Schmitz M, Vukicevic M, Laenen MV, Celi LA et al: Scalable Predictive Analysis in Critically Ill Patients Using a Visual Open Data Analysis Platform. *PLoS One* 2016;11:e0145791
48. Nepogodiev D, Martin J, Biccari B, Makupe A, Bhangu A, National Institute for Health Research Global Health Research Unit on Global Surgery: Global burden of postoperative death. *Lancet* 2019;393:401
49. Masino AJ, Harris MC, Forsyth D, Ostapenko S, Srinivasan L, Bonafide CP, et al: Machine learning models for early sepsis recognition in the neonatal intensive care unit using readily available electronic health record data. *PLoS One* 2019;14:e0212665
50. Aushev A, Ripoll VR, Vellido A, Aletti F, Pinto BB, Herpain A et al: Feature selection for the accurate prediction of septic and cardiogenic shock ICU mortality in the acute phase. *PLoS One* 2018;13:e0199089
51. Contreras I, Vehi J: Artificial Intelligence for Diabetes Management and Decision Support: Literature Review. *J Med Internet Res* 2018;20:e10775
52. Lundberg SM, Nair B, Vavilala MS, Horibe M, Eisses MJ, Adams T, et al: Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng* 2018;2:749–760
53. Cabitza F, Rasoini R, Gensini GF: Unintended Consequences of Machine Learning in Medicine. *JAMA* 2017; 318:517–518
54. Povyakalo AA, Alberdi E, Strigini L, Ayton P: How to discriminate between computer-aided and computer-hindered decisions: a case study in mammography. *Med Decis Making* 2013;33:98–107



55. Lapuschkin S, Wäldchen S, Binder A, Montavon G, Samek W, Müller KR: Unmasking Clever Hans predictors and assessing what machines really learn. *Nature Communications* 2019;10:1096. DOI: 10.1038/s41467-019-08987-4
56. Cannesson M, Shafer SL: All Boxes Are Black. *Anesth Analg* 2016;122:309–317
57. Tapan S, Wang D: A Further Study on Mining DNA Motifs Using Fuzzy Self-Organizing Maps. *IEEE Trans Neural Netw Learn Syst* 2016;27:113–124
58. Koza W, Filippo D, Cotik V, Stricker V, Munoz M, Godoy N, et al: Automatic Detection of Negated Findings in Radiological Reports for Spanish Language: Methodology Based on Lexicon-Grammatical Information Processing. *J Digit Imaging* 2019;32:19–29
59. Rodriguez A, Smielewski P, Rosenthal E, Moberg D: Medical Device Connectivity Challenges Outline the Technical

Requirements and Standards For Promoting Big Data Research and Personalized Medicine in Neurocritical Care.

*Mil Med* 2018;183:99–104

60. Roger VL, Boerwinkle E, Crapo JD, Douglas PS, Epstein JA, Granger CB, et al: Strategic transformation of population studies: recommendations of the working group on epidemiology and population sciences from the National Heart, Lung, and Blood Advisory Council and Board of External Experts. *Am J Epidemiol* 2015;181:363–68
61. Hawkins BE, Huie JR, Almeida C, Chen J, Ferguson AR: Data Dissemination: Shortening the Long Tail of Traumatic Brain Injury Dark Data. *J Neurotrauma* 2019 (Online First). DOI: 10.1089/neu.2018.6192
62. Hoorweg AJ, Pasma W, van Wolfswinkel L, de Graaff JC: Incidence

of Artifacts and Deviating Values in Research Data Obtained from an Anesthesia Information Management System in Children. *Anesthesiology* 2018;128:293–304.

### Korrespondenz- adresse



**Prof. Dr. med.  
Bettina Jungwirth**

Klinik für Anästhesiologie  
Universitätsklinikum Ulm  
Albert-Einstein-Allee 23  
89081 Ulm, Deutschland

E-Mail: [bettina.jungwirth@uni-ulm.de](mailto:bettina.jungwirth@uni-ulm.de)

ORCID-ID: 0000-0001-9749-7460