

Künstliche Intelligenz im ärztlichen Alltag

Einsatzgebiete in der Medizin:

Nutzen, Herausforderungen und Forderungen der FMH



Herausgeberin

FMH – Verbindung der Schweizer
Ärztinnen und Ärzte, Bern

Textbearbeitung

hallokern.ch

Konzept, Layout und Illustrationen

atelierrichner.ch

Publikation

September 2022

www.fmh.ch | ehealth@fmh.ch

Farbkennzeichnung Abbildungen

Die grafischen Darstellungen in dieser Broschüre sind in den Farben Ocker und Blau gehalten.

- Ocker verweist auf Bezüge zum Menschen.
- Blau verweist auf Bezüge zur Maschine.

Inhalt

Abkürzungen	4
5 Vorwort	
6 Welchen Nutzen erwarten Ärztinnen und Ärzte von künstlicher Intelligenz?	
8 Was ist künstliche Intelligenz?	
9 Methoden der künstlichen Intelligenz	
16 In welchen Bereichen der Medizin kann künstliche Intelligenz zur Anwendung kommen?	
20 Vor welche Herausforderungen stellt künstliche Intelligenz die Medizin?	
1. Fairness	20
2. Autonomie und Kontrolle	21
3. Transparenz	24
4. Verlässlichkeit	25
5. Sicherheit	27
6. Datenschutz	27
Zwischenfazit: Welcher menschliche Akteur muss dafür Sorge tragen, dass ein KI-System vertrauenswürdig ist und bleibt?	29
30 Wie wird künstliche Intelligenz in der Medizin reguliert?	
32 Welche Forderungen stellt die FMH an die künstliche Intelligenz in der Medizin?	
35 Wie wird sich das Berufsbild Ärztin/Arzt in Zukunft wandeln?	
40 Literatur	

Abkürzungen

ADR	Adverse Drug Reaction
AMA	American Medical Association
BGE	Schweizerisches Bundesgericht
CanMEDS	Canadian Medical Education Directives for Specialists
CDSS	Clinical Decision Support Systems
CNN	Convolutional Neural Network
CONSORT-AI	Consolidated Standards of Reporting Trials-Artificial Intelligence
CPME	Standing Committee of European Doctors
CPRD	Clinical Practice Research Datalink
DDI	Drug-drug interaction
DNN	Deep Neural Network
DR	Diabetische Retinopathie
EbM	Evidenzbasierte Medizin
EHR	Electronic health record /elektronische Patientenakte
EKG	Elektrokardiogramm
E-Person	Elektronische Person
FDA	U.S. Food and Drug Administration
GAN	Generative Adversarial Network
GPU	Graphical Processing Unit
HLEG	High-Level Expert Group on AI
ICMRA	International Coalition of Medicines Regulatory Authorities/ Internationale Koalition der Arzneimittelbehörden
IP	Intellectual Property
IRGC	International Risk Governance Center
IVDR	Regulation on in-vitro diagnostic medical devices/ Verordnung (EU) 2017/746 des Europäischen Parlaments und des Rates vom 5. April 2017 über In-vitro-Diagnostika
KG	Krankengeschichte
KI	Künstliche Intelligenz
MDR	Medical Device Regulation
MepV	Medizinprodukteverordnung 812.213
MI-CLAIM	Minimum information about clinical artificial intelligence modelling
ML	Machine Learning
MRI	Magnetresonanztomographie
NDD	Neural network-based method for drug-drug interaction prediction
NLP	Natural Language Programming
NN	Neural Network
PPG	Photoplethysmographie
RF	Random Forest
RNN	Recurrent Neural Network
SAMW	Schweizerische Akademie der medizinischen Wissenschaften
SATW	Schweizerische Akademie der technischen Wissenschaften
SPIRIT-AI	Standard Protocol Items: Recommendations for Interventional Trials-Artificial Intelligence
SVM	Support Vector Machine
TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis

Vorwort

Künstliche Intelligenz (KI) scheint allgegenwärtig zu sein, so oft wie wir diesem Begriff begegnen. Es reicht ein Blick in die Zeitung und schon ist die Rede vom autonomen Fahren, der Überwachung von Bürgerinnen und Bürgern oder vom Einsatz der KI zur Bewältigung von juristischen Klagen und Entscheiden.

Wir begegnen nicht nur dem Begriff, wir verwenden KI auch täglich. Ein fremdsprachiger Text, ja sogar ein Gesetzestext, wird in wenigen Sekunden in die gewünschte Sprache übersetzt. Wir haben dadurch barrierefreien Zugang zur Berichterstattung dieser Welt über alle Landes- und Sprachgrenzen hinweg. Auch die gesprochene Sprache kann in Bruchteilen von Sekunden in einen maschinenlesbaren Text umgewandelt und inhaltlich erschlossen werden. Wir fahren, der Stimme unseres «intelligenten» Telefons folgend, zum Zielort unserer Wahl, ohne je einen Blick auf eine Karte werfen zu müssen. Selbst Staus sind kein Hindernis mehr, werden wir doch einfach daran vorbei geleitet.

Aber wer von uns fuhr noch nie im Auto entlang einer holprigen, sehr schmalen Forststrasse, welche eher den Namen Forstweg verdient hätte, eben dieser netten Stimme folgend, mit dem mulmigen Gefühl es handle sich hier wohl um eine eher «ungewöhnliche» Nebenroute. Eine solche Wegleitung ist an sich harmlos, solange sie nicht in einem Abgrund oder in einem See endet. Ihr unkritisch zu folgen, könnte hingegen in der Medizin fatale Auswirkungen haben. Unnötige Umwege oder gar Schaden müssen auf alle Fälle vermieden werden.

Machen wir uns also gemeinsam auf die Suche nach den Antworten auf folgende Fragen:

- Was verbirgt sich hinter dem Begriff «künstliche Intelligenz»?
- Was kann sie leisten und was nicht?
- Welches sind die grossen Herausforderungen in der Anwendung von KI im medizinischen Bereich?
- Wie können wir deren Risiken minimieren und wie deren Chancen maximieren?
- Welche Forderungen an den Einsatz von KI leiten wir als Ärztinnen und Ärzte daraus ab?

Wenn man sich mit der künstlichen Intelligenz eingehender beschäftigt, merkt man, dass sie zwar kein heilsbringendes 8. Weltwunder ist, aber auch kein Golem, der – erst einmal ausser Kontrolle geraten – zu einer Gefahr seines Schöpfers wird. Wir wollen mit dieser Broschüre KI sozusagen entmystifizieren und Ihnen helfen, sich über die Möglichkeiten und Grenzen von KI im medizinischen Bereich ein realistisches Bild machen zu können. Letztlich und hauptsächlich zum Vorteil unserer Patientinnen und Patienten.

Ich wünsche Ihnen viel Spass bei der Lektüre!



Dr. med. Alexander Zimmer,
Mitglied des FMH-Zentralvorstandes,
Departementsverantwortlicher
Digitalisierung/eHealth

Welchen Nutzen erwarten Ärztinnen und Ärzte von künstlicher Intelligenz?

Zweifelloos bewirkt künstliche Intelligenz in allen Bereichen der Medizin einen tiefgreifenden Wandel, der bereits im Gange ist. Die Ärztinnen und Ärzte in der Schweiz sind jedoch dezidiert der Ansicht, dass die persönliche Beziehung zwischen Arzt und Patient auch in einer von künstlichen Intelligenz geprägten Welt im Mittelpunkt stehen muss.

Im Jahr 1950 verdoppelte sich das medizinische Wissen noch alle 50 Jahre, 2010 war die Zeitspanne auf 3,5 Jahre geschrumpft, und 2020 fand die Verdoppelung gemäss Schätzungen alle 73 Tage statt¹. Die Digitalisierung im Gesundheitsbereich generiert mehr und mehr digital verfügbare Daten. Diese können prinzipiell in die Diagnostik und in den Behandlungsentscheid einbezogen werden, um die Diagnose- und Behandlungsqualität zu verbessern. Die neusten Computergenerationen verfügen über eine enorme Rechenleistung und können mit Hilfe von grossen, multimodalen Datensätzen krankheitsrelevante Hypothesen testen. Computer sind im Vergleich zu Menschen in der Lage, Informationen systematischer zu verarbeiten, Entscheidungen konsistenter zu treffen und schneller auf Veränderungen zu reagieren².

Die Fachliteratur beschreibt drei Ebenen, auf die sich KI in der Medizin auswirkt:

1. auf Ärztinnen und Ärzte in Form von computer-gestützten Erkennungs- und Entscheidungssystemen, z.B. durch schnelle, genaue Bildinterpretation;
2. auf Gesundheitssysteme durch die Verbesserung von Arbeitsabläufen und durch Minimierung des medizinischen Fehlerpotenzials, aber auch durch die Anpassung der Versorgung, beispielsweise hin zu KI-basiertem Monitoring in Kombination mit Telemedizin von zu Hause aus;
3. auf Patientinnen und Patienten, indem diese ihre eigenen Daten zur Förderung ihrer Gesundheit verarbeiten können, und durch die Veränderung der Beziehung zwischen Ärztin und Patientin^{3,4}.

Ärztinnen und Ärzte sehen den Nutzen von künstlicher Intelligenz primär in technischen und methodischen Optimierungen entlang der Patient Journey.

Die Resultate der von der FMH durchgeführten Umfrage «Digital Trends Survey 2021» zum Nutzen digitaler Gesundheitsanwendungen aus Sicht der Ärzteschaft und Bevölkerung verweisen hingegen eindrücklich auf die Grenzen der Akzeptanz von KI: Ärztinnen und Ärzte sehen den Nutzen solcher Anwendungen in der Vereinfachung von administrativen Prozessen beziehungsweise Arbeitsabläufen. Sie stellen sich gegen Anwendungen, die den Arztbesuch sukzessive ersetzen sollen. So sind sie überzeugt, dass der Faktor Mensch für den Behandlungserfolg wichtig und durch KI nicht zu ersetzen ist. Beispielsweise lehnen 94% der Ärzteschaft eine ausschliesslich durch eine intelligente Software gestellte Diagnose ab. 85% wollen keine rein softwaregetriebene Erstellung des Behandlungsplanes. Aus der Sicht der befragten Ärztinnen und Ärzte wie auch der an der Studie beteiligten Patientinnen und Patienten soll die Digitalisierung vor allem mehr Zeit für die persönliche Behandlung durch die Ärztin oder den Arzt schaffen⁵.

Für die Ärzteschaft sehen wir einen möglichen Nutzen von KI in der Verbesserung und Weiterentwicklung entlang der Patient Journey in folgenden Bereichen:

-
- **Prävention / Prognose** Mit KI lassen sich neue biologische Mechanismen einer Krankheit, Risikofaktoren oder protektive Faktoren identifizieren.
-
- **Diagnostik** Klinische Entscheidungsunterstützungssysteme (Clinical Decision Support Systems, CDSS) ermöglichen beispielsweise eine schnelle, genaue Befundinterpretation. Sie können die Variabilität der Entscheidung minimieren und die Konsistenz und Effizienz der Entscheidung verbessern.
-
- **Behandlung** Der Einsatz von CDSS kann durch Erkennung der Nebenwirkungen und der Arzneimittelwechselwirkung zu einer Erhöhung der Arzneimittelsicherheit führen und die Überwachung verbessern. Mit KI-Systemen liessen sich die Patientenidentifizierung und-rekrutierung für klinische Studien vereinfachen, indem sie elektronische medizinische Patientenakten (electronic health records, EHR) nach geeigneten Patientinnen und Patienten durchsuchen.
-
- **Patientenmonitoring** KI kann die Überwachung des Behandlungsfortschritts und der Medikamenteneinhaltung vereinfachen. So liessen sich unkomplizierte, persönliche ärztliche Konsultationen teilweise durch virtuelle und telemedizinische Konsultationen ersetzen.
-
- **Verwaltung** KI reduziert den Aufwand von sich wiederholenden administrativen Arbeitsabläufen, wie die Eingabe von Patientendaten (z.B. Extraktion von EHR) und die automatisierte Überprüfung von Labordaten und Bildgebungsergebnissen. So würde sie Ressourcen beim Gesundheitsfachpersonal freisetzen, das sich verstärkt auf die zwischenmenschlichen Aspekte der Medizin konzentrieren könnte.
-

Was ist künstliche Intelligenz?

Neu ist KI nicht: Bereits vor weit über 80 Jahren war die Idee von der denkenden Maschine in der Wissenschaft ein Thema. Ihre aktuelle Bedeutung erlangte KI mit den rasant zunehmenden Leistungen der Computer.

Bis heute gibt es keine allgemeingültige Definition von KI⁶. Der englische Mathematiker Alan Turing war wahrscheinlich der erste, der sich mit intelligenten Maschinen beschäftigte. Bereits 1947 formulierte er auf einem Symposium in Manchester die zentrale Frage zu KI: «Können Maschinen denken?» John McCarthy, einer der Gründerväter von KI und ihr Namensgeber, definiert sie als «Wissenschaft und Technik der Herstellung intelligenter Maschinen, insbesondere intelligenter Computerprogramme»⁷. Gebräuchlich ist auch die Definition von KI als «die Lehre von Berechnungen, die das Wahrnehmen, Denken und Handeln ermöglichen»⁸.

Neu ist KI also nicht. Warum ist sie in den letzten Jahren über viele Bereiche des menschlichen Lebens hinweg zu einem omnipräsenten Thema geworden? Ihren fulminanten Aufschwung verdankt sie der deutlich verbesserten Rechenleistung und Speichertechnologien, die das Verarbeiten grosser Datenmengen ermöglichen. Das Schlagwort dazu lautet Big Data und bedeutet für den Medizinalsektor z. B. hochauflösende, medizinische Bildgebung, Genomsequenzierung, Biosensoren aber auch elektronische medizinische Patientenakten.

Die Funktion von künstlicher Intelligenz gliedert sich im Wesentlichen in die Aktionen Wahrnehmung, Analyse und Entwicklung von Handlungsoptionen.

Beschreiben lässt sich KI beziehungsweise die intelligente Maschine als computerbasiertes System (KI-System), das für einen vom Menschen definierten Zweck Vorhersagen, Empfehlungen oder Entscheidungen treffen kann, und so reale Umgebungen beeinflusst. KI verwendet maschinelle und/oder menschliche Eingaben, um:

1. reale Umgebungen wahrzunehmen;
2. solche Wahrnehmungen automatisiert in Bezug auf ein gewisses Ziel oder Ergebnis zu analysieren (z. B. mit Machine Learning, ML) und daraus Modelle abzuleiten;
3. Schlüsse aus dem Modell zu ziehen (Modellinferenz), um so Erkenntnisse und Handlungsmöglichkeiten zusammenzustellen⁶.

Zudem sind KI-Systeme so konzipiert, dass sie mit unterschiedlichem Grad an Autonomie arbeiten^{6,9}.

Ein anschauliches Beispiel für die drei genannten Charakteristika von KI ist das autonome Fahren:

1. Autonome Fahrzeuge verfügen über eine Vielzahl von Sensoren und Kameras, die z. B. mit Ultraschall oder Funkwellen arbeiten und so ihre Umgebung wahrnehmen.
2. Mit Hilfe von Deep Learning, einem Teilbereich des ML, d. h. einer der KI zugrunde liegenden Methodik, werden diese Bilder und Sensorimpulse analysiert und mit dem Ziel «unfallfreies Fahren» in ein Modell für ein eigenständiges (autonomes) Fahren umgewandelt, d. h.
3. das KI-System trifft mithilfe des Modells und den während des Fahrens erhobenen Daten die Entscheidung, wann es fährt, bremst und anhält. Der Fahrer kann im Notfall eingreifen und das KI-System des autonomen Fahrzeugs übersteuern¹⁰.

Nachdem das Wesen von KI in den Grundzügen aufgezeigt ist, stellt sich die Frage der Methodik, die ihr zugrunde liegt. Welche Methoden gibt es, und worin unterscheiden sie sich?

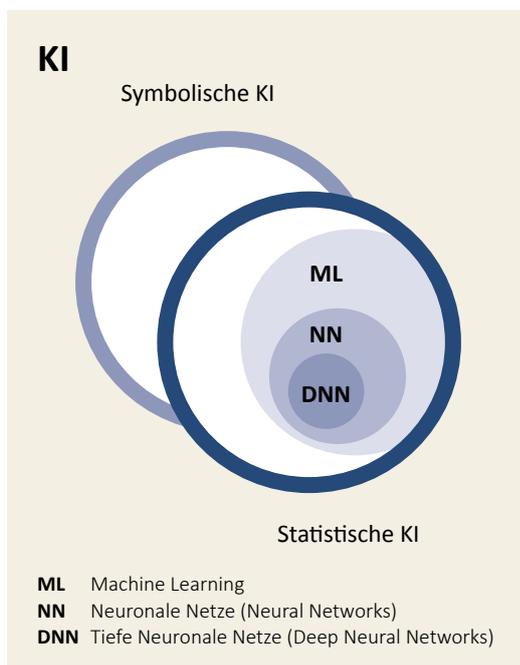
Methoden der künstlichen Intelligenz

In der jüngeren Geschichte der künstlichen Intelligenz verfolgte die Forschung zwei grundverschiedene Stossrichtungen: Den symbolischen und den statistischen Ansatz. Heute liegt das Schwergewicht auf statistischer künstlicher Intelligenz, bei der komplexe Probleme mit mathematischen Modellen und statistischen Verfahren gelöst werden.

Historisch hat die KI-Forschung symbolische KI von statistischer KI unterschieden⁶ (Abb. 1). Symbolische KI dominierte die KI-Forschung von Mitte der 1950er bis Ende der 1980er Jahre. Ziel war es, herauszufinden, ob beziehungsweise inwiefern menschliche Intelligenz auf die Darstellung von Symbolen reduziert werden kann. Die KI-Forscher gingen von der Hypothese aus, dass jedes intelligente System – ob Mensch oder Maschine – funktionieren muss, indem es Datenstrukturen verarbeitet, die aus Symbolen bestehen¹¹. Um symbolische KI auf den Weg zu bringen, galt es, detaillierte Entscheidungsfindungsstrukturen zu entwickeln. Diese sollten die Komplexität der realen Welt übersetzen und Maschinen dabei helfen, menschenähnliche Entscheidungen zu treffen⁶. Zur symbolischen KI zählen beispielsweise Techniken wie die logische Programmierung (logic programming), die Beschreibungslogik (description logic), die symbolische Syntaxanalyse (symbolic parsing), und die Regelinduktion (rule induction)¹².

Ab den 1990er Jahren gewann die statistische KI immer mehr an Bedeutung¹¹. Während symbolische KI menschliche Regeln abbildet, versucht das statistische KI-System Muster in Daten zu erkennen und daraus ein Modell abzuleiten⁶. Statistische KI umfasst beispielsweise Techniken wie Bayes'sche Netze, verborgene Markov-Modelle, Markov-Entscheidungsprozesse, statistisches Parsing und neuronale Netze¹². Es gibt auch KI-Systeme, die auf einer Kombination des symbolischen und des statistischen Ansatzes beruhen. Ein Beispiel ist die maschinelle Sprachverarbeitung (Natural Language Programming, NLP). Diese stützt sich meist auf grosse Datenmengen (statistischer Ansatz) und berücksichtigt gleichzeitig die Regeln der Grammatik (symbolischer Ansatz)⁶.

Abb. 1 **Methodische Ansätze von künstlicher Intelligenz im Überblick**



Im Bereich der statistischen künstlichen Intelligenz sind die Technologien des Machine Learning angesiedelt. Dabei lernt der Computer – ohne explizite Programmierung – aus Daten Vorhersagen zu treffen.

Zur statistischen KI gehört das Machine Learning (ML). Unter diesem Begriff wird eine Reihe von Methoden zusammengefasst, die es Computern ermöglichen, ohne explizite Programmierung aus Daten zu lernen und Vorhersagen zu treffen oder zu verbessern. Machine Learning verkörpert einen Paradigmenwechsel. Es macht den Schritt von der «normalen Programmierung», bei der alle Anweisungen explizit an den Computer gegeben werden

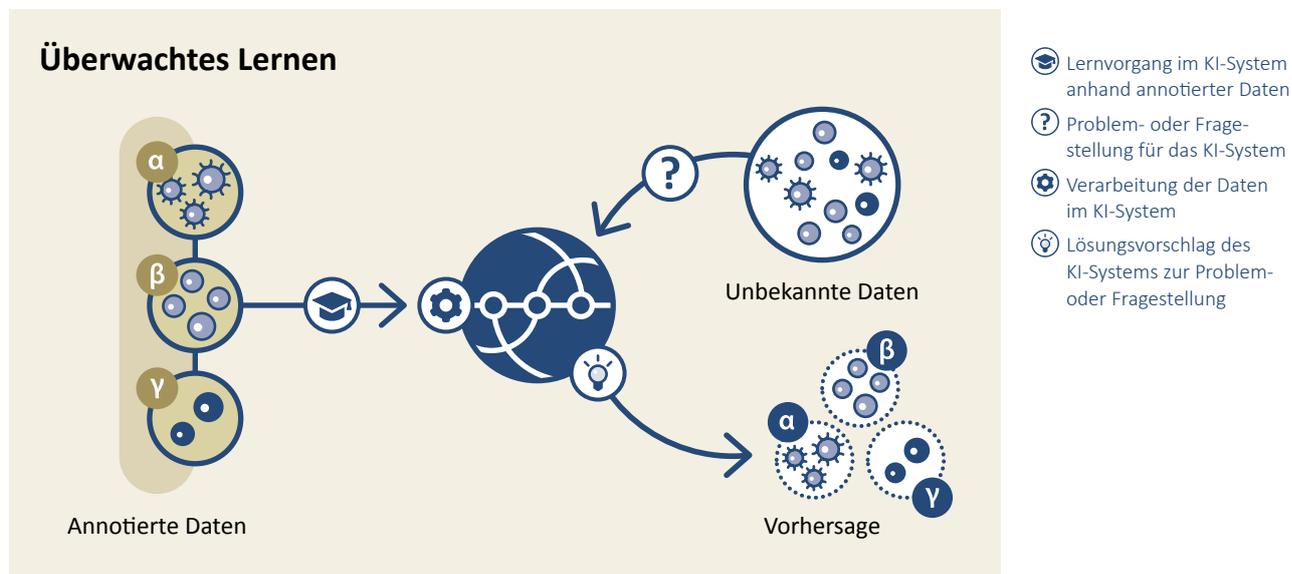
müssen, hin zur «indirekten Programmierung», die auf Grundlage von Daten Vorhersagen trifft¹³. Computer, die Methoden des Machine Learning verwenden, sind fähig, sich eigenes Wissen anzueignen, indem sie beispielsweise in Rohdaten Muster erkennen und auf deren Basis Vorhersagen beziehungsweise Schlüsse aus komplexen Daten ableiten¹⁴. Zu den im Machine Learning verwendeten Techniken gehören beispielsweise die lineare und logistische Regression, Entscheidungsbäume, die Hauptkomponentenanalyse (principle component analysis), Neuronale Netze (Neural Network, NN) und Tiefe Neuronale Netze (Deep Neural Network, DNN)⁶.

Derzeit gibt es im ML drei verschiedene Ansätze wie diese Systeme lernen (Abb. 2a–c):

- überwachtes Lernen
- unüberwachtes Lernen
- bestärkendes Lernen

KI-Systeme können aus mehreren unterschiedlichen KI-Methoden aufgebaut sein. Z. B. lassen sich Machine-Learning-Algorithmen als Bausteine eines KI-Systems verwenden. Dies befähigt das KI-System dazu, komplexe Probleme zu lösen. Es unterteilt diese in «einfachere» Problemstellungen, die jeweils von einem Machine-Learning-Algorithmus bearbeitet werden¹⁵.

Abb. 2a Machine-Learning-Lernansatz überwachtes Lernen *Supervised learning*



Charakteristik

Überwachtes Lernen fasst die Gruppe der KI- Algorithmen^{a)} zusammen, die einen annotierten Datensatz benötigen, um zu lernen. Die Annotation erfolgt typischerweise durch Hinzufügen von weiteren Informationen wie Metadaten (z. B. Diagnosecodes), die einen Datensatz – d. h. medizinische Bilder, Texte oder andere Daten – inhaltlich beschreiben. Dieser Datensatz muss somit sogenannte Eingabe-Ausgabe Paare enthalten: z. B. digitale Bilder einer pathologischen Hautläsion und die exakte kodierte Diagnose dazu.

^{a)} Ein Algorithmus ist ein Verfahren für die Lösung eines mathematischen Problems durch eine endliche Anzahl von Schritten. Einfach ausgedrückt ist der Algorithmus ein Regelwerk, dem ein Computer oder eine Maschine folgt, um ein bestimmtes Ziel zu erreichen²³.

«Überwacht» bedeutet, dass das maschinelle Lernen mit Daten erfolgt, die vorher von einem Menschen annotiert werden mussten¹⁶. Unter Verwendung des annotierten Datensatzes wird der Algorithmus trainiert. Er lernt, die Abhängigkeiten zwischen Ein- und Ausgabe zu identifizieren und ist nun in der Lage, aufgrund des erkannten Musters auf unbekanntem, nicht annotierten Datensätzen Vorhersagen zu treffen¹⁴. Beim überwachten Lernen ist im Voraus bekannt, welche Art von Ergebnis erwartet werden kann¹⁷.

Verwendete Daten

Annotierter Datensatz mit Ein- und Ausgabepaaren Zugrundeliegende Daten können bereits strukturiert (z. B. Laborwerte) oder unstrukturiert (z. B. Bilder) sein.

Verwendete Techniken

- Lineare und logistische Regression
- Support Vector Machine (SVM)
- Naive Bayes-Klassifikation (naive Bayes)
- Random Forest (RF)¹⁹
- Entscheidungsbäume^{17,18}
- Neuronale Netze (NN) (nicht abschliessend)

Nutzen

Klassifizierung Daten werden einer bestimmten Kategorie zugeordnet.

Regression Daten werden einem Zahlenwert zugeordnet.

Anwendungsgebiete

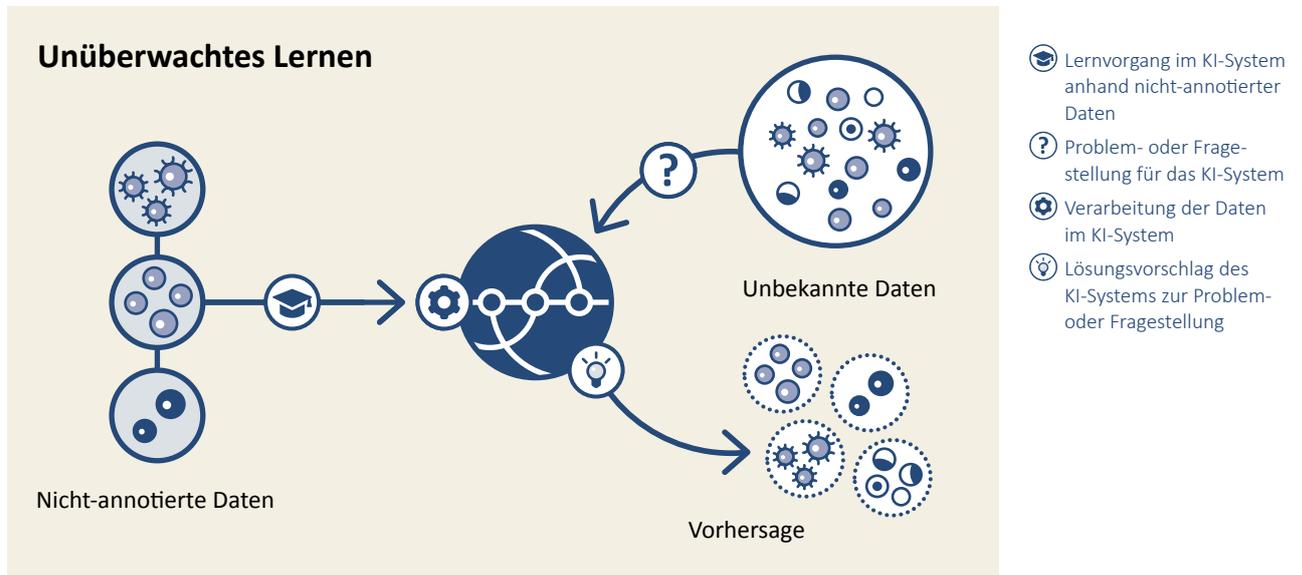
Klassifizierung Medizinische Diagnose, Gesichtserkennung¹⁶

Regression Dosierung von Medikamenten, Vorhersage des Todeszeitpunkts²⁰, Temperaturvorhersage¹⁶

Grenzen

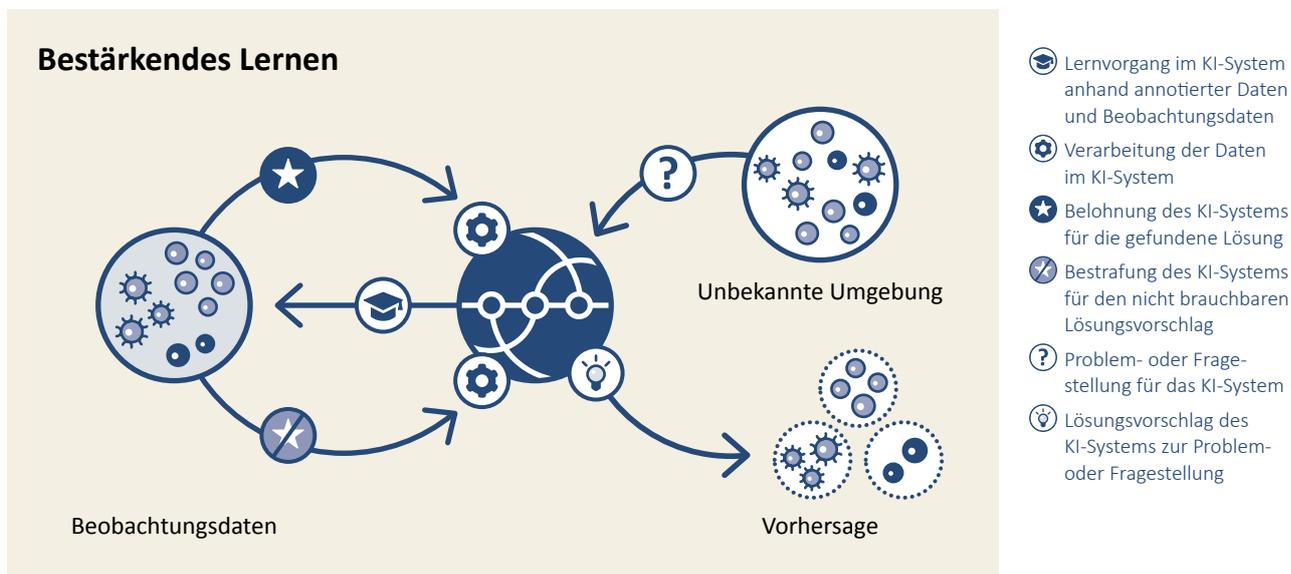
Benötigt grosse annotierte Datensätze für Training und Validierung, idealerweise mit einer unabhängigen Kohorte¹⁹.

Abb. 2b **Machine-Learning-Lernansatz unüberwachtes Lernen** *Unsupervised learning*



Charakteristik	Unüberwachtes Lernen fasst die Gruppe der KI-Algorithmen zusammen, die versuchen, aus nicht-annotierten Daten Muster zu erkennen oder Schlüsse zu ziehen. Im unüberwachten Lernen gibt es, anders als im überwachten Lernen, keine «richtigen» Antworten ¹⁶ . Unüberwachtes Lernen zielt darauf ab, aus grossen Mengen von Daten neue Erkenntnisse zu gewinnen. Der Algorithmus des unüberwachten Lernens bestimmt selber, was innerhalb des Datensatzes anders oder interessant ist ¹⁷ .
Verwendete Daten	<i>Nicht-annotierter Datensatz</i> Daten können strukturiert (z. B. Laborwerte) oder unstrukturiert (z. B. Bilder) sein.
Verwendete Techniken	<ul style="list-style-type: none"> – Probabilistische Clustering Verfahren – K-Means Clustering¹⁸ – Hierarchische Clusteranalyse²¹ <p>(nicht abschliessend)</p>
Nutzen	<i>Mustererkennung und Datenclustering</i> Ähnliche Daten werden unterschiedlichen Gruppen (Cluster) zugeordnet ¹⁶ .
Anwendungsgebiete	<p><i>Mustererkennung und Datenclustering</i></p> <ul style="list-style-type: none"> – Erkennen von unterschiedlichen biologischen Mechanismen eines ansonsten «homogenen» Krankheitsbildes – Identifizierung versteckter Muster in Genotyp- oder Phänotypdaten^{17,19} <p><i>Reduzierung der Datendimensionalität</i> Komplexe Datensätze werden durch deren Clusterzugehörigkeit repräsentiert: Beschleunigung anderer ML-Algorithmen, indem die Anzahl der Berechnungen reduziert wird^{16,17}.</p>
Grenzen	Die Validierung der Cluster ist schwierig. Sie kann z. B. durch Messung der Stabilität der Cluster in unabhängigen Kohorten ¹⁹ erfolgen oder durch Annotation und Messung der Clusterzugehörigkeit ähnlicher Datenelemente ²¹ .

Abb. 2c **Machine-Learning-Lernansatz bestärkendes Lernen** *Reinforcement learning*



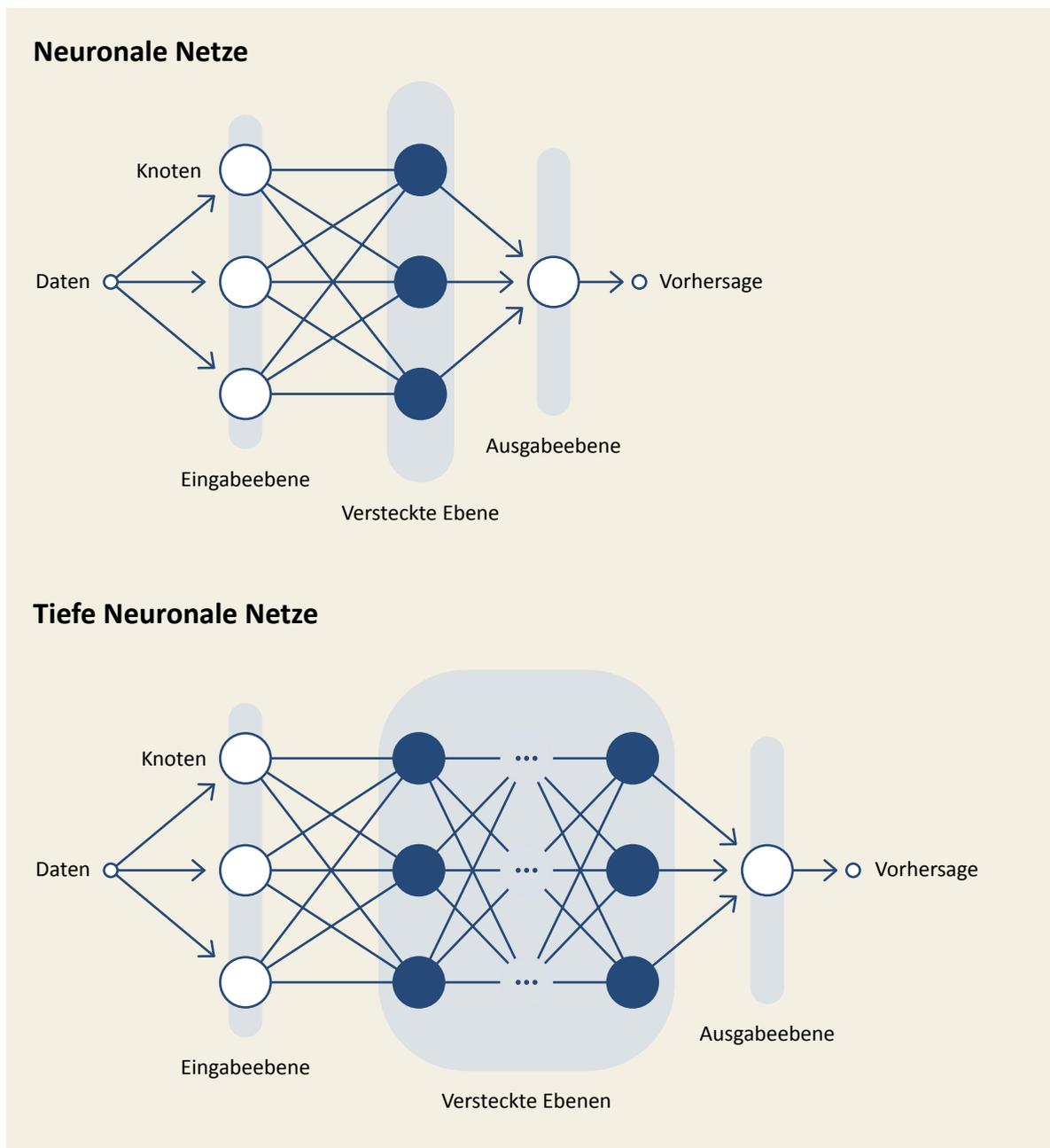
Charakteristik	Bestärkendes Lernen fasst die Gruppe der KI-Algorithmen zusammen, die durch Interaktion mit der Umgebung lernen. Der Algorithmus probiert eigenständig Lösungswege aus und erhält Rückmeldung aus seiner Umgebung. D. h., er wird für den Lösungsweg entweder belohnt oder bestraft. Durch die wiederholte Handlung und Rückmeldung lernt er, welche Auswirkung welche Handlung in welcher Situation hat ¹⁴ .
Verwendete Daten	Annotierte Daten Diese dienen typischerweise als Grundlage, die dem KI-Algorithmus initial zum Trainieren zur Verfügung stehen. In einer zweiten Phase generiert der KI-Algorithmus zusätzliche Daten aus der Interaktion mit seiner Umgebung (Beobachtungsdaten) ¹⁵ .
Verwendete Techniken	Identisch mit den Techniken des überwachten Lernens. Das Trainieren des Algorithmus erfolgt über eine Belohnungsmaximierung: Der Algorithmus lernt, die optimalen Entscheidungen zu treffen ¹⁵ .
Nutzen	Erlernen wie die Umgebung funktioniert ¹¹ .
Anwendungsgebiete	<ul style="list-style-type: none"> – Behandlungsstrategien (z. B. optimale Behandlungsstrategie von Sepsis auf der Intensivstation ²²) – Robotische Steuerung – KI in Spielen ¹⁴
Grenzen	Speicherintensiv ¹⁹ Die Optimierung kann speziell im medizinischen Bereich aus Sicherheitsgründen nicht durch die direkte Interaktion mit der Umgebung (z. B. dem Arzt) erfolgen ²² .

Deep Learning ist Machine Learning unter Einbezug von tiefen Neuronalen Netzen

Mit Machine Learning verbunden ist das Deep Learning. Dieses findet seit 2012 immer mehr Beachtung^{15,24}. Deep Learning basiert auf der Technik der Tiefen Neuronalen Netze (Deep Neuronal Networks DNN), die wiederum aus den Neuronalen Netzen (NN) hervorgegangen sind (Abb. 3).

Neuronale Netze gibt es bereits seit den 1950er Jahren²⁵. Ab 2006 wurden sie zu Tiefen Neuronalen Netzen weitentwickelt: «Networks went deep»^{15,26}. Eine Grundvoraussetzung dafür war, dass zu diesem Zeitpunkt wesentlich grössere Datensätze («Big Data») zur Verfügung standen, und dass die Hardware der Computer, respektive die Prozessoren wesentlich effizienter und schneller geworden waren¹⁵.

Abb. 3 Eigenschaften des Neuronalen Netzes und des Tiefen Neuronalen Netzes



Art des Netzes	Charakteristik
Neuronales Netz (NN)	<p>Neuronale Netze bestehen aus einer Eingabeebene (input layer), <u>einer versteckten Ebene (hidden layer)</u> und einer Ausgabebene (output layer). Jede dieser Ebenen ist aus Knoten (nodes) aufgebaut. Ein Neuronales Netz kann Millionen von Knoten enthalten.</p> <hr/> <p>Die Knoten der verschiedenen Ebenen erinnern an Neuronen des menschlichen Gehirns^{27,28}. Jeder Knoten ist gewichtet, besitzt einen Schwellenwert und ist mit jedem Knoten der nächsten Ebene verbunden. Wenn das Ausgabesignal (Output) eines Knotens über dessen Schwellenwert liegt, sendet dieser Knoten seine Information an die Knoten der nächsten Ebene: Die Computerarchitektur lernt durch Versuch und Irrtum²⁴.</p>
Tiefes Neuronales Netz (DNN)	<p>Tiefe Neuronale Netze unterscheiden sich von Neuronalen Netzen in der Anzahl von Ebenen. Neuronale Netze umfassen einschliesslich Ein- und Ausgabebene nur drei Ebenen. Demgegenüber gliedern sich Tiefe Neuronale Netze in die Eingabeebene (input layer), <u>zwei bis mehrere versteckte Ebenen (hidden layers)</u> und die Ausgabebene (output layer)²⁸.</p> <hr/> <p>Tiefe Neuronale Netze können Milliarden von Knoten in Dutzenden von Ebenen enthalten.</p> <hr/> <p>Varianten des Tiefen Neuronalen Netzes sind Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), sowie Generative Adversarial Networks (GAN).</p> <ul style="list-style-type: none"> – Convolutional Neural Networks kommen bei der Bilderkennung, Personenidentifikation, autonomer Robotik, selbstfahrenden Autos und in der Spracherkennung zum Einsatz. – Recurrent Neural Networks finden Verwendung bei der Spracherkennung und für Übersetzungen. – Generative Adversarial Networks eignen sich beispielsweise zum Generieren von neuen Bildern und Fotos. Auch das Phänomen «Deep Fakes» basiert auf dieser Technik¹⁴.

Der Unterschied zwischen Deep Learning und klassischem Machine Learning

Klassische Machine-Learning-Algorithmen verwenden für ihre Vorhersagen typischerweise strukturierte Datensätze. Möglich ist auch die Verwendung von unstrukturierten Daten, sofern diese entsprechend vorbereitet sind (pre-processing). D.h. sie müssen manuell in ein strukturiertes Format gebracht werden.

Beim Deep Learning entfällt ein Teil dieser manuellen Datenvorbereitung. Deep-Learning-Algorithmen können mit unstrukturierten Daten (z. B. Bilder, Text) arbeiten, ohne dass diese zuvor strukturiert wurden. Die Algorithmen des Deep Learning extrahieren selbständig die wesentlichen Merkmale aus den unstrukturierten Daten damit sie sich voneinander unterscheiden²⁹. In der Medizin kommt Deep Learning beispielsweise zum Einsatz, um in medizinischen Scans, Pathologiepräparaten, Bildern von Hautläsionen, Netzhautbildern, Elektrokardiogrammen und Vitaldaten Muster zu erkennen und diese zu interpretieren³.

In welchen Bereichen der Medizin kann künstliche Intelligenz zur Anwendung kommen?

Es ist vorstellbar, dass Anwendungen von künstlicher Intelligenz den Menschen von seiner Entstehung bis zum Tod in allen Belangen seiner Gesundheit begleiten. Die in diesem Kapitel dargestellten konkreten Anwendungsbeispiele aus jedem Bereich der Patient Journey vermitteln ein Bild von den Einsatzmöglichkeiten künstlicher Intelligenz in der Medizin.

KI im Bereich Prävention / Prognose am Beispiel Vorhofflimmern

Vorhofflimmern ist eine Herzrhythmusstörung, die mit erhöhter Morbidität und Mortalität einher geht. Diese Indikation ist der Hauptrisikofaktor für einen kardioembolischen Schlaganfall und entscheidend für dessen Prävention. Durch unmittelbar am Körper getragene Geräte, sogenannte Wearables (z. B. Smartwatch, Smartphone), kann der Herzrhythmus kontinuierlich überwacht und Vorhofflimmern frühzeitig entdeckt werden. Ermöglicht wird dies durch optische Sensoren, die Schwankungen des Blutvolumens im mikrovaskulären Gewebebett aufzeichnen.

Diese Technologie nennt sich Photoplethysmographie (PPG). Sie benötigt eine Lichtquelle und einen Photodetektor. Das damit generierte Signal besteht aus zwei Komponenten: Eine Gleichstromkomponente repräsentiert das reflektierte respektive transmittierte Licht des sich im Ruhezustand befindenden arteriellen Bluts, des venösen Bluts, der Haut und des Gewebes. Eine pulsierende Wechselstromkomponente entsteht durch die Veränderung der Lichtabsorption aufgrund von Änderungen des arteriellen Blutvolumens. Da Bewegung Artefakte in den generierten Signalen erzeugen kann, sind die tragbaren Geräte mit Beschleunigungssensoren ausgestattet, die die Beschleunigungskräfte entlang verschiedener Raumrichtungen messen.

Mit Hilfe der PPG lassen sich physiologische Parameter wie Herzfrequenz, Blutdruck, Herzleistung, Sauerstoffsättigung ableiten³⁰. Das tragbare Gerät «lernt» die Herzfrequenz der Trägerin oder des Trägers im Ruhezustand sowie unter physischer Aktivität. Wenn das kontinuierlich gemessene Signal signifikant abweicht, schlägt das Gerät Alarm und empfiehlt beispielsweise ein EKG. Dieses kann gegebenenfalls mit dem gleichen tragbaren Gerät erstellt werden³. Die Methoden, die sich diese Algorithmen bedienen, reichen von statistischen Ansätzen (z. B. logistische Regression, Markov Modell) über Machine Learning (z. B. Support Vector

Machine, Entscheidungsbäume, K-nearest neighbors) bis hin zu Deep Learning (z. B. Convolutional Neural Network, CNN)³⁰.

KI im Bereich Diagnostik am Beispiel diabetischer Retinopathie

Diabetische Retinopathie (DR) kann zu einer erheblichen Verringerung der Sehkraft und zu Erblindung führen, wenn sie nicht frühzeitig erkannt und behandelt wird³¹. Mit Hilfe der Fundusphotographie werden Bilder der Netzhaut, des Sehnervkopfes und der Makula aufgenommen. Die Bilder ermöglichen die Erkennung und Überwachung von Krankheiten wie DR, Glaukom, Neoplasien der Netzhaut und altersbedingte Makuladegeneration³². In den letzten Jahren hat sich die optische Kohärenztomographie als Referenzmethode für die Erkennung des diabetischen Makulaödems etabliert³¹.

In einer prospektiven Studie in zehn verschiedenen Kliniken der Primärversorgung wurden Patientinnen und Patienten mit Diabetes, aber ohne bekannte Retinopathie, mit Hilfe eines proprietären bildgebenden Gerätes untersucht. Das System verwendete Fundusphotographien und optische Kohärenztomographie der Netzhaut und interpretierte diese mit einem Deep Learning Algorithmus (Convolutional Neural Network, CNN). Die gemachten Aufnahmen wurden mit Fundus-Referenzdaten verglichen. Vor Studienbeginn absolvierten die Anwender des von KI unterstützten Geräts ein standardisiertes Trainingsprotokoll. Das Gerät wurde von der FDA für die Verwendung von Gesundheitsfachpersonen zur Diagnose des diabetischen Makulaödems und mittelschwerer bis schwerer DR zugelassen^{31, 33}.

Anwendungen von künstlicher Intelligenz gibt es in fast allen Bereichen der medizinischen Fachdisziplinen. Auch bei extrem heiklen chirurgischen Eingriffen kann künstliche Intelligenz Unterstützung leisten.

KI im Bereich Behandlung am Beispiel Wechselwirkungen von Arzneimitteln

Die Bezeichnung Arzneimittelwechselwirkungen (drug-drug interaction, DDI) steht für unerwünschte Nebenwirkungen, die durch die gleichzeitige Einnahme von zwei oder mehr Medikamenten entstehen. Diese gegenseitige Beeinflussung der Arzneimittel kann weitere Erkrankungen hervorrufen und im schlimmsten Fall zum Tod führen (Adverse Drug Reaction, ADR)³⁴. Betroffen sind hier oftmals ältere Patientinnen und Patienten, die generell eine erhöhte Erkrankungsprävalenz haben und eher multimorbid sind³⁵.

Mit Hilfe von Deep Learning lassen sich neue Arzneimittelwechselwirkungen vorhersagen. Ein Beispiel für eine solche Prognose ist die Methode NDD (Neural network-based method for drug-drug interaction prediction). NDD ist ein mehrstufiger Prozess, bei dem Informationen über chemische Medikamenteneigenschaften und Medikamentennebenwirkungen in einem Neuronalem Netz zur Vorhersage von neuen Arzneimittelwechselwirkungen verwertet werden³⁴.

KI im Bereich Patientenmonitoring am Beispiel Antizipation eines zukünftigen Krankheitsverlaufs

KI kann das Patientenmonitoring in verschiedenen Krankheitsstadien unterstützen. Dabei greift sie auf longitudinale Daten zurück. Das sind Daten, die über einen gewissen Zeitraum erhoben wurden (z.B. mit Sensoren) und eine Aussage über den zukünftigen Krankheitsverlauf zulassen. Ein gutes Beispiel sind multimodale Sensordaten in der Intensivmedizin, über die KI eine Aussage über ein anstehendes Kreislaufversagen machen kann³⁶. Ein weiteres Beispiel ist die Aufzeichnung von EKG – Signalen mit Hilfe von Wearables zur Vorhersage einer Verschlechterung der Herzinsuffizienz³⁷.

KI im Bereich Verwaltung am Beispiel Dokumentation

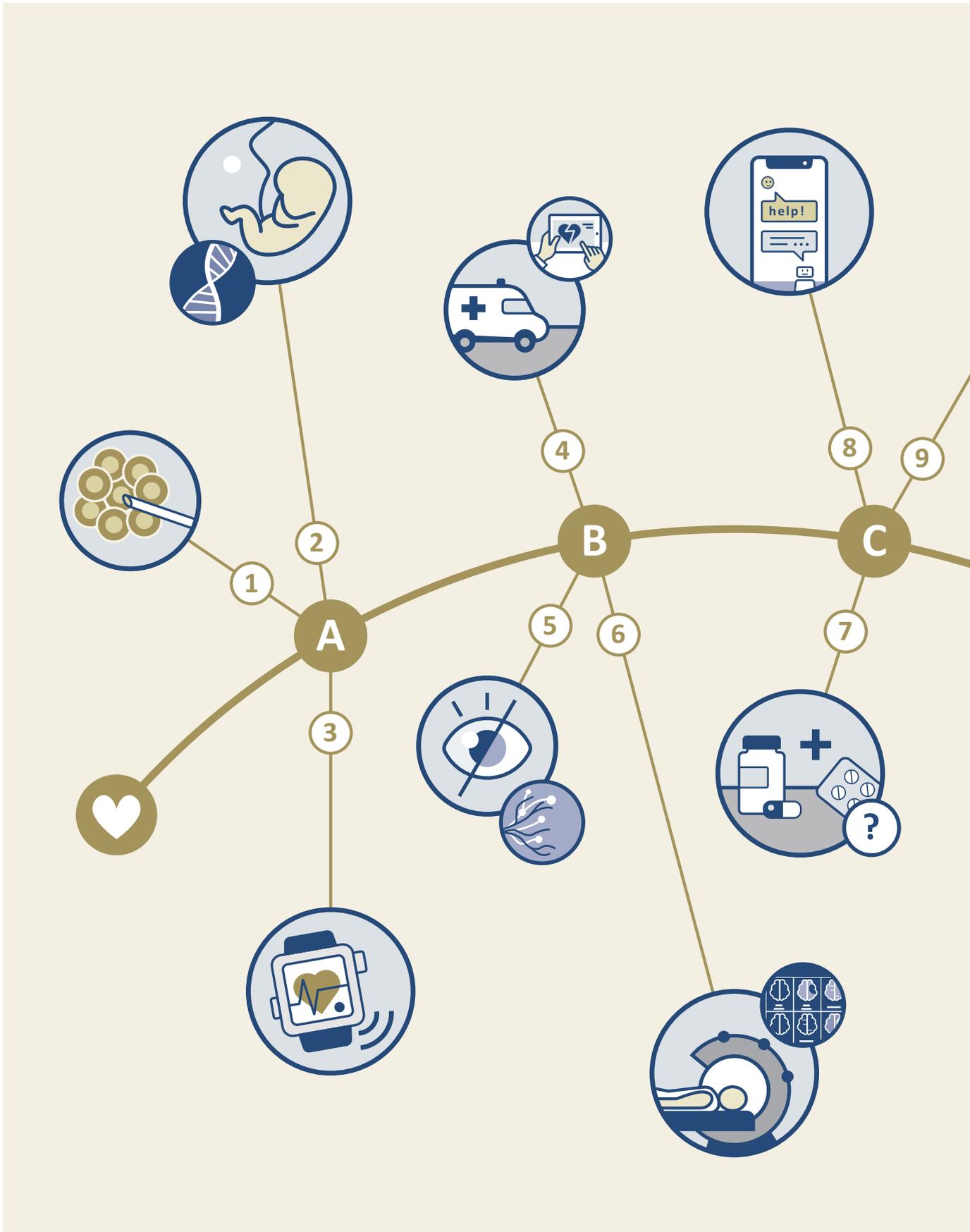
Die medizinische Dokumentation entlang des Patientenpfads ist fester Bestandteil der täglichen Routine von Gesundheitsfachpersonen. Schweizer Ärztinnen und Ärzte in der Akkutsomatik verbringen täglich 20% ihrer Arbeitszeit für Dokumentationsarbeiten rund um das Patientendossier. Bei Assistentenärztinnen und Assistentenärzten in der Akkutsomatik sind es sogar 27%. Damit wenden diese für die Dokumentation nahezu gleich viel Zeit auf wie für patientennahe Tätigkeiten³⁸.

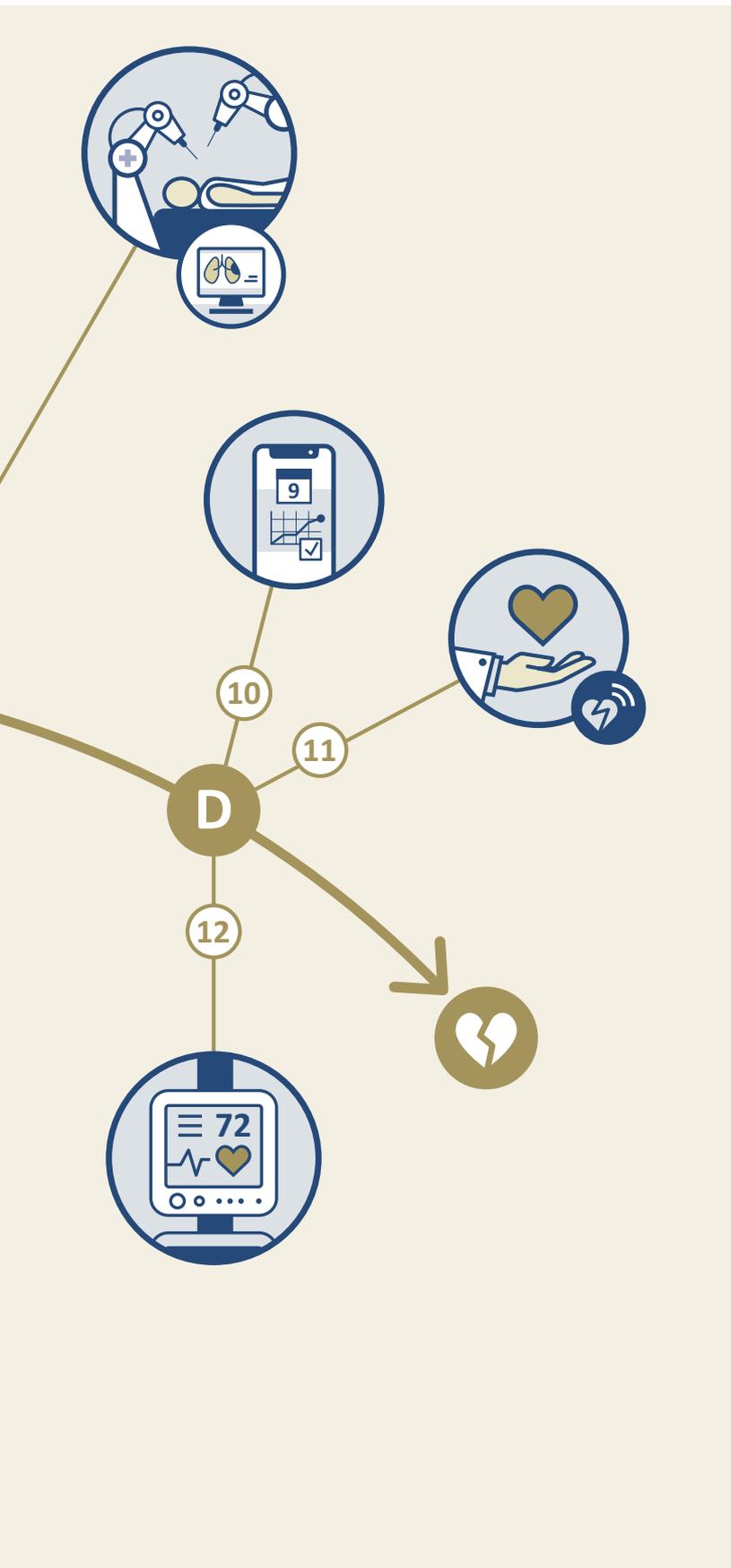
Die Verwendung der elektronischen Patientenakte (EHR) erhöht die Verfügbarkeit von medizinischen Daten³⁹. Allerdings kann eine wenig anwenderfreundliche EHR auch zu erhöhten Dokumentationszeiten führen. Dies auf Kosten der effektiven Zeit für die Betreuung der Patienten und der Zufriedenheit bei der Ärzteschaft. Um Dokumentationszeiten und Fehler in der medizinischen Dokumentation zu reduzieren, stehen Ärzteteams in den USA «klinische Schreiber» (Scribes) zur Verfügung. Scribes sind für die Zusammenarbeit mit Klinikern geschult. Sie dokumentieren medizinische Informationen innerhalb ärztlicher Konsultationen aussagekräftig und genau und ermöglichen somit den Ärztinnen und Ärzten im Team, sich besser auf die medizinischen Aspekte der Konsultation zu konzentrieren. Eingesetzt werden jetzt auch «Digitale Schreiber». Diese nutzen Fortschritte in der Spracherkennung und der Verarbeitung von natürlicher Sprache mit KI (Natural Language Processing, NLP), um Ärztinnen und Ärzten die Dokumentation zu erleichtern⁴⁰. Die Spracherkennung erlaubt es den Ärzten, Notizen zu diktieren oder Gespräche mit Patientinnen und Patienten automatisch zu dokumentieren⁴¹. Dabei werden Gespräche nicht nur Wort für Wort transkribiert, sondern bei Bedarf auch zusammengefasst. Darüber hinaus kann KI auch verständliche klinische Berichte direkt von Primärdaten (Röntgen- oder Pathologiebildern) generieren⁴².

Breit gefächertes Anwendungsspektrum

Anwendungen von künstlicher Intelligenz gibt es in fast allen Bereichen der medizinischen Fachdisziplinen, beispielsweise der Radiologie, Pathologie, Dermatologie, Ophthalmologie, Kardiologie, Gastroenterologie, Chirurgie, Psychiatrie-Psychotherapie. Auch extrem heikle mikrochirurgische Eingriffe wie z.B. im Inneren des Auges⁴³, können mit Hilfe von KI-unterstützten robotischen Systemen durchgeführt werden.

Abb. 4 Mögliche Anwendungen von künstlicher Intelligenz entlang des menschlichen Lebens





A Prävention und Prognostik

1. Embryo Selektion bei In-vitro Befruchtung
2. Genomanalyse und Genominterpretation bei Neugeborenen
3. Herzfrequenzmessung via Smartwatch

B Diagnose

4. Diagnose eines Herzinfarktes oder Schlaganfalls
5. Prävention von Erblindung bei Diabetischer Retinopathie
6. Auswertung von medizinischer Bildgebung

C Behandlung

7. Identifikation von unerwünschten Wechselwirkungen zwischen verschiedenen Arzneimitteln
8. Unterstützende Behandlung von Depression oder Angstzuständen über Chatbots oder Apps
9. Extrem präzise Ausführung von Operationen durch Roboter

D Patientenmonitoring/Nachsorge

10. Tägliche Check-ups, z. B. nach schwerer Krankheit, und Dokumentation des Therapiefortschritts mittels Apps
11. Verbesserung der Patientensicherheit: Warnung an Ärztinnen und Pflegepersonal, wenn Messgrößen ausserhalb des Normbereiches liegen
12. Vorhersage des Todeszeitpunkts im Krankenhaus mit dem Ziel, die Massnahmen von Palliative Care entsprechend auszurichten

Vor welche Herausforderungen stellt künstliche Intelligenz die Medizin?

Der Einsatz von künstlicher Intelligenz in der Medizin eröffnet neue Möglichkeiten, bringt aber auch Herausforderungen, respektive Risiken mit sich⁴⁴. Diese lassen sich anhand der sechs Dimensionen der Vertrauenswürdigkeit beschreiben: Fairness, Autonomie und Kontrolle, Transparenz, Verlässlichkeit, Sicherheit und Datenschutz.

Die sechs Dimensionen der Vertrauenswürdigkeit orientieren sich an den Kernanforderungen für vertrauenswürdige KI der europäischen High-Level Expert Group on AI (HLEG)⁴⁵ und basieren auf dem Fraunhofer KI-Prüfkatalog⁴⁶.

1. Fairness

Das Prinzip Fairness verbietet ungerechtfertigte Ungleichbehandlung und damit unzulässige Diskriminierung⁴⁷. Jegliche menschliche Tätigkeit wird auch durch ungewollte und unbewusste Befangenheit (Verzerrung/Bias) beeinflusst³⁹. KI-Systeme lernen aus vorhandenen historischen Daten (z. B. medizinische Daten), die unter Umständen solche Verzerrungen enthalten. Deshalb sind auch KI-Systeme nicht vor Verzerrungen gefeit.

Solche Verzerrungen können zu medizinischen Fehlern mit Folgen für Patientinnen und Patienten und/oder zu sozioökonomischem Ungleichgewicht in der Gesundheitsversorgung führen. Verwendet ein KI-System beispielsweise Daten aus einer unvollständigen elektronischen Patientenakte, können bei geschützten («protected») Patientengruppen Verzerrungen durch fehlende Daten (missing data bias), durch zu kleine Stichprobengrößen (minority bias) und durch Messfehler beziehungsweise Fehlklassifikationen (informativeness bias) entstehen^{48, 49}.

Vielfältige Ursachen für Verzerrungen

Belastbare Vorhersagen kann ein KI-System nur für eine Patientenpopulation mit X Eigenschaften treffen, für die es trainiert, getestet und validiert wurde. Wird das KI-System bei einer Patientenpopulation eingesetzt, auf die es nicht trainiert, getestet und validiert wurde, kann es zu Fehlvorhersagen kommen (Training-serving skew)⁴⁹.

Verzerrungen können nicht nur durch die für das Trainieren verwendeten Datensätze entstehen, sondern auch durch das Design des KI-Systems sowie durch die Interaktion des KI-Systems mit Ärztinnen und Ärzten auf der einen und Patientinnen und Patienten auf der anderen Seite.

Designbedingte Verzerrungen

Im Design des KI-System können Verzerrungen durch die Verwendung von wenig aussagekräftigen Variablen entstehen. Dies ist der Fall, wenn eine Variable statt einer medizinischen Wahrheit eher Ungleichheiten in der Gesundheitsversorgung beschreibt (label bias). Eine weitere durch das KI-Design verursachte Verzerrung entsteht, wenn Kohorten verwendet werden, die nur leicht zu messende Patientenpopulationen enthalten und weitere, für die medizinische Generalisierbarkeit der zu treffenden Vorhersagen wichtige Patientenpopulationen auslassen (cohort bias: z. B. Bias in der Geschlechterverteilung)⁴⁹.

Interaktionsbedingte Verzerrungen

Verzerrungen durch die Interaktion des KI-Systems mit der Ärzteschaft entstehen durch Unwissenheit, z. B.:

- wenn der Arzt nicht weiss, dass das KI-System für eine bestimmte Patientengruppe ungenauere Vorhersagen macht (automation bias⁵⁰), oder
- wenn die Ärztin die Empfehlungen des KI-Systems übernimmt, obwohl dies in diesem bestimmten Fall nicht angezeigt ist und das KI-System aus diesen fehlerhaften Entscheiden weiter lernt (feedback loop bias)⁴⁹, oder
- wenn der Arzt ungewollt oder gewollt Warnhinweise des KI-Systems für geschützte Patientengruppen ignoriert (dismissal bias)⁴⁹.

Patientenbedingte Verzerrungen

Auf den Patienten zurückzuführende Verzerrungen im KI-System können auch dann entstehen, wenn KI beispielsweise bei einem Teil der Patientenpopulation aufgrund technologischer oder sozialer Hürden gar nicht angewandt werden kann (privilege bias). Patientengruppen, die schlechte Erfahrungen mit der Technologie gemacht haben, können dieser grundsätzlich misstrauen. Sie wollen sich deshalb auf diese Art nicht behandeln lassen und erhalten somit unter Umständen keine angemessene Gesundheitsversorgung (informed mistrust). Gleiches geschieht, wenn eine bestimmte Patientengruppe keine «Stimme» in der Entwicklung, Verwendung und Bewertung der KI-Systeme erhält (agency bias)⁴⁹.

Nützliche Testverfahren

Mit Hilfe von Tools wie FairTest können Vorhersagen von KI-Algorithmen und deren Assoziationen mit sensiblen Eigenschaften wie Geschlecht, Ethnie usw. auf Fairness getestet werden. FairTest enthält auch Debugging Funktionen, die es Programmierern ermöglichen, Auslöser für die identifizierten unfairen Effekte zu beseitigen^{51, 52}.

2. Autonomie und Kontrolle

Der Begriff Autonomie bezieht sich in diesem Kontext auf zwei Aspekte: einerseits die Selbstbestimmung des Menschen (z. B. der Nutzerin eines KI-Systems), andererseits die Autonomie beziehungsweise den Handlungsspielraum des KI-Systems. Gemeint ist, dass für das jeweilige KI-System die angemessene und verantwortungsvolle Aufgabenteilung zwischen Mensch/Nutzer und KI-System definiert werden muss^{46, 47}.

Folgende Szenarien sind für die Zusammenarbeit zwischen Mensch und Maschine denkbar (Abb. 5):

«In the Loop»

Der Mensch ist für den Arbeitsprozess unabdingbar. Das bedeutet, dass die Maschine keine Aufgabe ohne die Bestätigung durch den Menschen abschliessen kann. Zudem kann der Mensch die vom KI-System selbständig getroffenen Entscheidungen überschreiben und kompensieren^{46, 53}.

«On the Loop»

Maschinen erledigen den Grossteil der Arbeit. Der Mensch ist unter Normalbedingungen in nur wenige oder keine Entscheidungen des KI-Systems einbezogen. Er wird zur Kontrollinstanz, die sicherstellt, dass der Arbeitsprozess normal und fehlerfrei abläuft. Er kann die vom KI-System selbständig getroffenen Entscheidungen nachträglich korrigieren, überschreiben und kompensieren^{46, 53}.

«Out of the Loop»

Menschliches Eingreifen ist nicht erforderlich. Die Maschinen sind präzise und autonom genug, um die Arbeit allein zu bewerkstelligen. Der Mensch kann nur noch entscheiden, ob er das KI-System einsetzen möchte oder nicht^{46, 53}.

Abb. 5 Denkbare Arten der Zusammenarbeit zwischen Mensch und Maschine

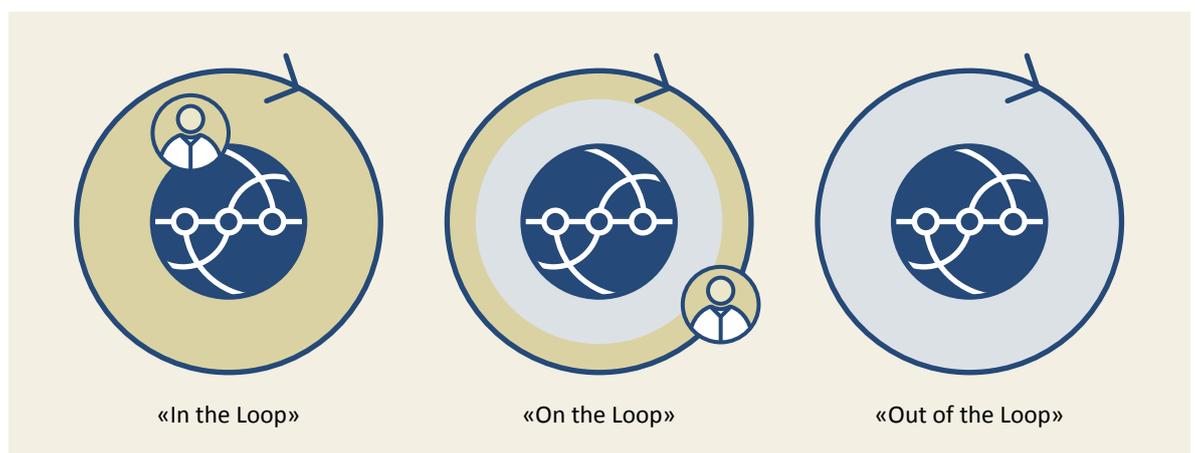
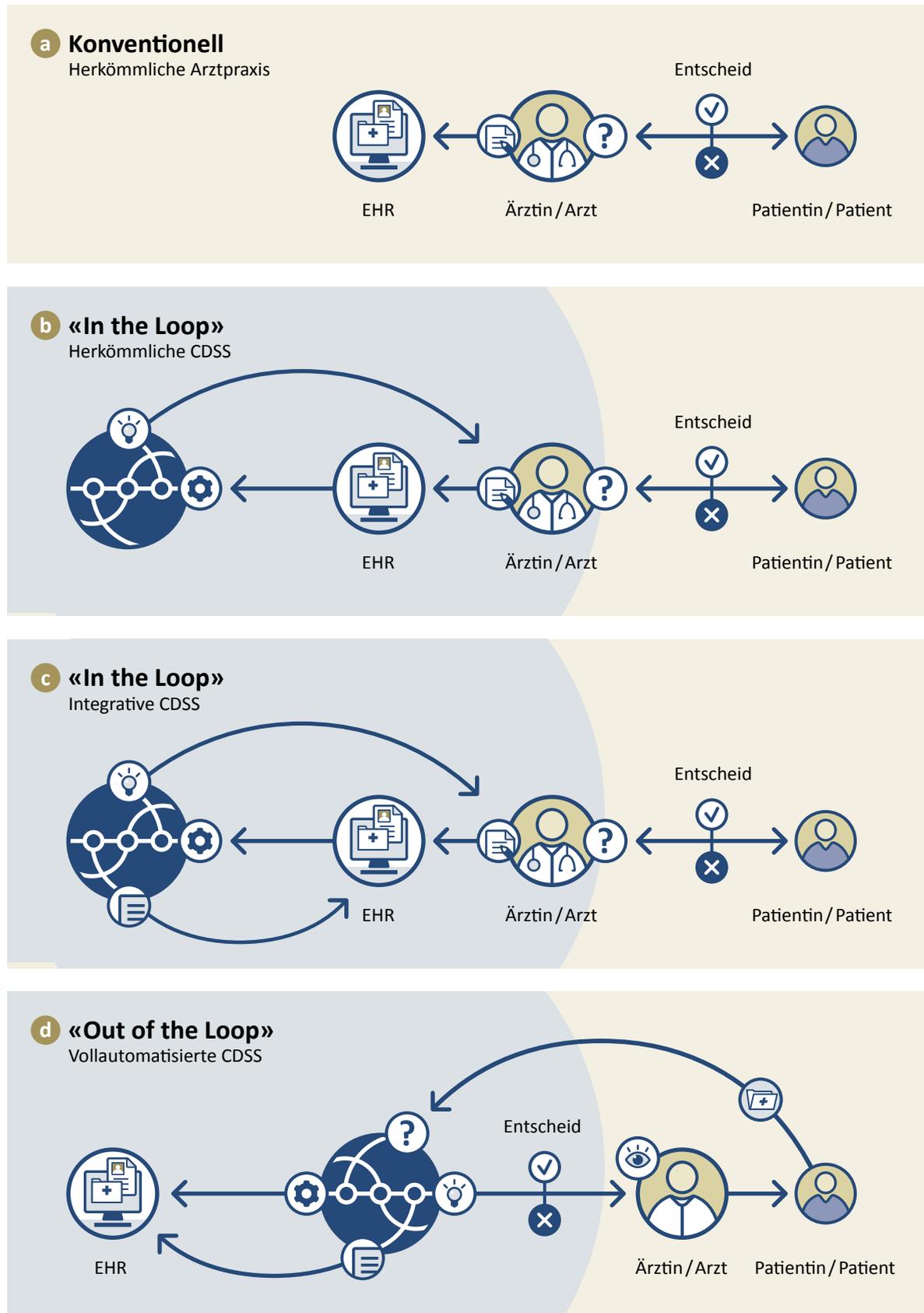


Abb. 6 Denkbare Zusammenarbeit von «In the Loop» und «Out of the Loop» CDSS mit Ärztinnen und Ärzten



- ① Problem- oder Fragestellung für die Ärztin/den Arzt (6a–c) oder das KI-System (6d)
- 📄 Dokumentation der Diagnose und Behandlung durch Ärztin/Arzt
- ⚙️ Verarbeitung der Daten im KI-System
- 💡 Lösungsvorschlag des KI-Systems zur Problem- oder Fragestellung
- 📄 Dokumentation Lösungsvorschlag durch KI-System
- 📄 Sammeln der medizinischen Patienteninformationen durch KI-System
- 👁️ Überprüfung des KI-Lösungsvorschlags durch Ärztin/Arzt
- ✅/❌ Genehmigung/ Ablehnung des KI-Lösungsvorschlags durch Ärztin/Arzt (6a–c) bzw. durch KI-System (6d)

Die Zusammenarbeit «In the Loop» und «Out of the Loop» zwischen Mensch und Maschine lässt sich auf die Anwendungen von KI in der Medizin folgendermassen übertragen (hier am Beispiel CDSS, Abb. 6):

Konventionell (Abb. 6a)

In einer herkömmlichen Arztpraxis erfasst die Ärztin oder der Arzt medizinische Informationen der Patientin, trifft medizinische Entscheidungen im Zusammenwirken mit der Patientin und dokumentiert Diagnose und Behandlung mit Hilfe der Patientenakte.

«In the Loop» (Abb. 6b und 6c)

Herkömmliche CDSS sammeln Informationen aus der elektronischen Patientenakte und unterbreiten der Ärztin oder dem Arzt mit Hilfe von regelbasierten oder Machine-Learning-Algorithmen Vorschläge, beispielsweise für die Diagnose und/oder den Behandlungsplan. Die finale Entscheidung trifft die Ärztin oder der Arzt (6b)³².

Bei integrativen CDSS kann der KI-Algorithmus selbständig auf die elektronische Patientenakte zugreifen, den Ärztinnen und Ärzten Vorschläge machen und diese selbständig in der elektronischen Patientenakte dokumentieren. Die Ärztin oder der Arzt trifft hier ebenfalls die finale Entscheidung (6c)³².

«Out of the Loop» (Abb. 6d)

In vollautomatisierten CDSS sammelt der KI-Algorithmus medizinische Informationen der Patientin oder des Patienten, trifft autonom medizinische Entscheidungen und dokumentiert diese in der elektronischen Patientenakte³².

Zunehmender Einsatz von «On the Loop»-Systemen

Noch existieren wenige «Out of the Loop»-Systeme ohne ärztliche Supervision, wie etwa solche zur automatischen Diagnose der diabetischen Retinopathie³¹. Grund sind die vielen technischen und regulatorischen Hürden, die den breiten Einsatz komplett autonomer KI-Systeme bremsen. Hingegen ist zu erwarten, dass sich «On the Loop»-Systeme zunehmend durchsetzen werden. Z. B. geht aus einer Studie hervor, dass KI die Pathologen bei der Diagnose des Prostatakarzinoms soweit unterstützen kann, dass diese nur noch 25 % der positiven Fälle beurteilen müssen. 75 % der negativen Fälle werden vom KI-System automatisch mit hoher Präzision aussortiert⁵⁴.

Zudem werden anhand von kommerziell erhältlichen KI-Systemen neue Arten von Interaktionen zwischen Arzt und KI sichtbar, welche die ärztliche Autonomie und Kontrolle ebenfalls betreffen können⁵⁵. Während z. B. die meisten angebotenen KI-Systeme des Typs «In the Loop» die Ärztinnen und Ärzte direkt unterstützen, gibt es verwandte Systeme, die eher indirekt im Hintergrund Unterstützung leisten.

- Ein Beispiel für direkte Unterstützung ist ein KI-System, das Ärzten durch eine Markierung das Auffinden einer Fraktur im Röntgenbild erleichtert⁵⁶.
- Ein Beispiel für eine indirekte Unterstützung ist ein KI-System, das die noch zu befundenen Radiologiebilder nach Schweregrad (oder Leidensdruck der Patienten) vorsortiert. So werden z. B. Radiologiebilder mit automatisch erkanntem Pneumothorax dem Arzt prioritär zur Befundung zugestellt⁵⁷.

KI-Unterstützung für nicht ärztliches Personal

Andere KI-Systeme erlauben es, ärztliche Prozesse an Dritte abzugeben. Beispielhaft ist ein System, das nicht ärztliche Gesundheitsfachpersonen dabei unterstützt, komplexe Echokardiogramm-Untersuchungen selbstständig durchzuführen⁵⁸. Die dabei erhobenen Daten können den Kardiologen zur Befundung weitergeleitet werden. Für die Ärztinnen und Ärzte resultiert daraus ein Zeitersparnis, für das nichtärztliche Gesundheitsfachpersonal eine Erweiterung der Kompetenz.

Akzeptanz und KI-Know-how der Ärzteschaft unabdingbar

Für die Integration von KI-Systemen in den klinischen Alltag oder den Alltag einer Fach-, respektive Hausarztpraxis ist die Akzeptanz der Ärzteschaft essenziell. Die Ärztinnen und Ärzte müssen bestimmen, wie sie am effizientesten mit der intelligenten Maschine zusammenarbeiten können. Deshalb ist es wichtig, dass sie in die Entwicklung von KI-Systemen für medizinische Anwendungen eingebunden werden.

Zusätzlich zu der Aufgabenverteilung zwischen Mensch und KI-System muss definiert werden, welche Vorkenntnisse Nutzerinnen und Nutzer haben müssen, um das KI-System korrekt anzuwenden, zu kontrollieren und gegebenenfalls zu überwachen. Dazu gehört, dass Ärztinnen und Ärzte über den Zweck, Nutzen, die Handhabung (Verwendung, Kontrolle, Überwachung) sowie die Grenzen und Risiken des KI-Systems ausreichend informiert werden. Sie müssen die Befähigung erlangen, das KI-System ihrer Sorgfaltspflicht entsprechend einzusetzen. Dieses Know-how trägt beispielweise dazu bei, den oben beschriebenen Automation Bias (Dimension Fairness) zu vermeiden⁴⁶.

3. Transparenz

Der Begriff Transparenz ist relevant für die Akzeptanz von KI und beinhaltet Aspekte der Nachvollziehbarkeit, Erklärbarkeit, Interpretierbarkeit und Reproduzierbarkeit. Wenn ein Mensch die Funktionsweise und Entscheidungen eines KI-Systems vollständig oder teilweise verstehen kann, spricht man von Transparenz oder Nachvollziehbarkeit. Wird eine Erklärung für das Zustandekommen einer Vorhersage des KI-Systems geliefert, spricht man von Erklärbarkeit. Wird für die Vorhersage eines KI-Systems eine Methode verwendet, die transparent ist (z. B. regelbasierte Modelle), spricht man von der Interpretierbarkeit des Modells (im Gegensatz zu Black-Box Modellen)^{46, 47}.

Black-Box Modelle sind meist robuster als beispielsweise regelbasierte (interpretierbare) Modelle. Allerdings sind sie oft kaum interpretierbar⁴⁷. Diesen Aspekt diskutiert die Fachwelt besonders im Zusammenhang mit Deep Neuronal Networks (DNN) und fordert, dass erklärbar und nachvollziehbar sein muss, wie KI-Systeme zu ihren Vorhersagen gelangen. Für DNN ist dies eher schwierig³. Die Forschungsarbeit zu dieser Frage ist in vollem Gange⁵⁹.

Internationale Empfehlungen und Initiativen

Es gibt mehrere internationale Initiativen, die sich den Themen Transparenz, Erklärbarkeit, Nachvollziehbarkeit angenommen haben. Einige Beispiele:

- TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) publiziert Empfehlungen für die Dokumentation des Studiendesigns, der Entwicklung, Validierung und Aktualisierung von KI-basierten Modellen, die Vorhersagen zu Prognosen oder Diagnosen machen⁶⁰.
- Die MI-CLAIM (Minimum information about clinical artificial intelligence modeling) Checkliste soll die transparente Dokumentation von KI-Algorithmen in der Medizin ermöglichen. Dabei geht es einerseits darum, die klinische Wirksamkeit, einschliesslich Fairness und Verzerrung (Bias) bewerten zu können. Andererseits soll die MI-CLAIM Checkliste dazu beitragen, den technischen Designprozess von klinischen KI-Systemen schnell nachvollziehbar zu machen^{61, 62}.
- CONSORT-AI (Consolidated Standards of Reporting Trials-Artificial Intelligence) ist eine neue Berichterstattungsleitlinie für klinische Studien, die medizinische Interventionen mit einer KI-Komponente evaluieren⁶³.
- Mit SPIRIT-AI (Standard Protocol Items: Recommendations for Interventional Trials-Artificial Intelligence) wurde parallel dazu eine Leitlinie für klinische Studienprotokolle entwickelt:⁶⁴.

Letztlich gibt es ein Spannungsfeld zwischen der Robustheit/Verlässlichkeit des KI-Systems und der Nachvollziehbarkeit/Transparenz von dessen Funktionsweise. Klar ist, dass Vorhersagen der verwendeten KI-Systeme für Ärztinnen und Ärzte soweit transparent sein müssen, dass sie diese korrekt anwenden können. Dies entsprechend ihrer vorausgesetzten Qualifikation⁴⁶. Zudem ist zu verhindern, dass irrelevante Details im Zusammenhang mit der Dokumentation der Transparenz zu einer Überforderung führen⁴⁷.

Eine komplexe Situation bei der Haftungsfrage

Aspekte der Nachvollziehbarkeit und Reproduzierbarkeit von KI-basierten Vorhersagen sind auch für rechtliche Fragen relevant. Z. B. in Bezug auf die Haftung und die strafrechtliche Verantwortlichkeit. Während ein einzelner Arztfehler einem Patienten schaden kann, könnte ein KI-System, das für eine ganze Patientenpopulationen diagnostische oder behandlungsrelevante Vorhersagen trifft, bei einem

systematischen Fehler einer Vielzahl von Patientinnen und Patienten schaden. Deshalb sind die systematische Fehlerbehebung, Audits, umfangreiche Simulationen und Validierungen sowie prospektive Validierungen im klinischen Alltag und im Praxisalltag erforderlich. Ziel muss es sein, das Fehlerisiko eines KI-Systems in der klinischen Praxis sehr gering zu halten³.

Wer haftet, wenn durch den Einsatz eines KI-Systems Patientinnen und Patienten einen Schaden zugefügt wurde – Arzt, Spital oder Hersteller? Die Antwort lautet: Die Haftung bei der Anwendung von KI im Rahmen einer medizinischen Behandlung wird derzeit entsprechend den Verantwortungsbereichen des behandelnden Arztes und des Herstellers beziehungsweise des sonstigen Verantwortlichen des KI-Systems aufgeteilt^{65,66}. Der Arzt kann für diejenigen Handlungen und Unterlassungen haftbar gemacht werden, die er erkennen, beeinflussen und auf Grund seiner Fachkompetenz zu verantworten hat^b. Dazu gehört auch der Entscheid, KI-Systeme bei der Diagnose und Behandlung einzusetzen, respektive auf den Einsatz von KI bei einer Behandlung zu verzichten, obwohl diese zum medizinischen Standard gehört^c.

Laufende Diskussion über mögliche Anpassungen der Haftungsgrundlagen

Eine Ärztin oder ein Arzt kann sich auch durch eine Überschreitung der eigenen Fachkompetenz haftpflichtig machen (Übernahmeverschulden)⁶⁷. Die Sorgfaltspflichten der Ärzte im Zusammenhang mit KI können sich je nach Anwendungsfall unterscheiden und müssten entsprechend definiert werden. Die Haftung der Ärztin oder des Arztes beurteilt sich derzeit weiterhin nach den bisherigen Haftungsgrundlagen⁶⁸. Sollten für den Einsatz von KI in der Medizin neue Rechtsgrundlagen geschaffen werden, sind diese so zu gestalten, dass sie die mit der Anwendung von KI verbundenen Gegebenheiten berücksichtigen. Zusätzlich zu den Ansprüchen wegen Schadensersatzes oder Genugtuung ist die strafrechtliche Verantwortlichkeit zu prüfen, falls beim Einsatz von KI eine vorsätzliche oder fahrlässige Verletzung der Sorgfaltspflicht durch den Arzt zu einer Schädigung des Patienten geführt hat. Die verschiedenen Aspekte einer denkbaren Revision der Haftungsgrundlagen werden zurzeit auf internationaler Ebene diskutiert.

^b BGE 133 III 121; BGE 130 IV 7 E. 3.3.

^c Zum Unterlassen Vergleiche:

BGE 134 IV 175 E. 3.2.; BGE 130 IV 7 E. 3.3.

4. Verlässlichkeit

Der Begriff Verlässlichkeit umfasst Aspekte der Qualität des KI-Systems. Zu diesen zählen die Performanz, die Robustheit, das Abfangen von Fehlern und die Einschätzung der Unsicherheit des verwendeten KI-Modells. Gemeint ist, dass getestet werden muss, wie fehleranfällig das KI-System ist. Diese Überprüfung klärt beispielsweise die Korrektheit der Vorhersagen des KI-Systems,

- wenn dieses seinem Zweck und seinen Anforderungen entsprechend eingesetzt wird,
- wenn es störungsanfällige Daten (beispielsweise Daten, die im Grenzbereich des Anwendungsgebietes liegen) oder manipulierte Eingabedaten (beispielsweise adversarial attacks) zur Prüfung der Robustheit erhält⁴⁶.

Ein KI-System kann nur in dem Spektrum korrekte Vorhersagen treffen, in dem es trainiert wurde⁴⁶. In der Medizin gibt es oft unstrukturierte Daten, beispielsweise in Form von Freitext-Anmerkungen des Arztes in einer Patientenakte oder einer Datenbank³. Um KI-Systeme zu trainieren braucht es für die jeweilige Patientenpopulation repräsentative, diversifizierte Datensätze von hoher Qualität. Nur so können diese Systeme erlernen, die richtigen Schlüsse für ihre Vorhersagen zu ziehen³⁹.

KI-Systeme nur bedingt auf weitere Patientenpopulationen übertragbar

Ein KI-System, das auf eine bestimmte Patientenpopulation in einer Klinik trainiert, getestet und validiert wurde, kann nicht automatisch auf eine andere Patientenpopulation in einer anderen Klinik übertragen werden (Generalisierbarkeit oder Portabilität). Bevor ein KI-System für eine neue Patientenpopulation zur Anwendung kommen kann, muss ein Test Klarheit schaffen, ob es korrekte Vorhersagen für diese neue Patientenpopulation trifft oder nicht. Gegebenenfalls ist eine Anpassung erforderlich⁶⁹. In diesem Zusammenhang ist zu prüfen, ob bei einem Algorithmus ein Overfitting vorliegt: Ist beispielsweise die Stichprobe (z. B. Grösse der Patientenpopulation) für die zu beantwortende medizinische Fragestellung zu klein, besteht die Gefahr, dass der KI-Algorithmus (z. B. Machine Learning) Datensatz-spezifische, d. h., nicht allgemeingültige Muster erkennt und darauf seine Vorhersagen stützt³. Die Generalisierbarkeit des Algorithmus wird dadurch eingeschränkt.

Trainingsdaten als wichtiger Qualitätsfaktor

Ebenfalls entscheidend für das Trainieren eines verlässlichen KI-Algorithmus sind die Qualität und die Verteilung der Trainingsdaten. So werden typischerweise im Gesundheitssystem vornehmlich Daten von Patientinnen und Patienten erfasst, bislang hingegen nur wenige Daten von gesunden Menschen.

Es braucht grosse qualitative hochwertige, annotierte medizinische Datensätze um KI-Systeme zu trainieren, zu testen und zu validieren.

Um beispielsweise Gehirnfunktionen und deren Beeinträchtigung mit Hilfe der medizinischen Bildgebung wie MRI zu untersuchen, braucht es zusätzlich zu den Aufnahmen des Gehirns von Patientinnen und Patienten auch Bilder von «gesunden» Gehirnen als Referenzdatensatz. Diese lassen sich für das Trainieren und Testen der dafür verwendeten KI-Systeme einsetzen. Auch um potentiell pathogene Genvarianten zu identifizieren, sind Daten der gesunden Bevölkerung zur Häufigkeitsverteilung von Genvarianten essenziell⁷⁰. Darüber hinaus müssen je nach Zweck des KI-Systems Qualitätsmerkmale wie beispielsweise Sensitivität und Spezifität und deren zu erreichende Werte festgelegt werden^{3,71}. So kann eine höhere Rate an Falschpositiven je nach Zweck tolerierbar sein.

Validierung unter Verwendung von unabhängigen Datensätzen

Um KI-Systeme auf ihre Übertragbarkeit (Portabilität) auf neue Patientenpopulation beziehungsweise Generalisierbarkeit, Robustheit und Leistungsfähigkeit (Performanz) zu testen, ist eine Validierung mit Hilfe von unabhängigen Datensätzen erforderlich. Theoretisch kommen zu diesem Zweck – zusätzlich zu den Datensätzen von Kliniken und Arztpraxen – beispielsweise Kohorten oder Datensätze aus klinischen und epidemiologischen Registern – in Frage. Die Validierung mit unabhängigen Datensätzen mit ähnlichen Merkmalen wie die für das Trainieren und Testen verwendeten Datensätze, zeigt, wie reproduzierbar die Leistungsfähigkeit des Modells ist.

Unabhängige Datensätze unterscheiden sich beispielsweise durch Abweichungen bei den Patientenmerkmalen, aber auch durch potenzielle Verzerrungen (Bias) oder Störfaktoren. Sie können auch manipulierte Daten enthalten. Die Validierung zeigt, wie übertragbar und robust das Modell ist, d. h., es werden die Generalisierbarkeit und Robustheit bewertet. Voraussetzung dafür ist, dass die Risiken entlang des gesamten Entwicklungsprozesses des KI-Systems identifiziert und die einzelnen Schritte validiert werden. Beispielsweise muss validiert werden, wie die für das KI-System relevanten Daten gesammelt, annotiert, getestet, mit diesen trainiert und validiert wurden. Zu diesem Zweck ist die Anwendung der im vorhergehenden Abschnitt erwähnten internationalen Standards für KI-unterstützte Studiendesigns (z. B. ML – Analyseplan) sinnvoll⁷².

Leistungsfähigkeit ist nicht gleich klinische Wirksamkeit

Wichtig in diesem Zusammenhang ist auch die Klärung der Fragen, weshalb eine bestimmte KI-Methode verwendet wurde und keine andere, und wie der Vergleich zu einer bestmöglichen alternativen KI-Methode getestet wurde. Hinzu kommt, dass die Validierung der Leistungsfähigkeit eines KI-Systems nicht gleichbedeutend ist mit der klinischen Wirksamkeit (efficacy). Diese betrachtet nicht nur die Treffsicherheit (accuracy) eines Systems sondern auch dessen Einfluss auf den Behandlungserfolg (patient outcome)^{73,74}. Dies impliziert, dass ein KI-System nachweislich das klinische Outcome verbessern und deshalb im realen klinischen Umfeld validiert sein sollte³. Zu berücksichtigen ist nicht zuletzt die Wirtschaftlichkeit, also die Kosten-Nutzen-Rechnung und die Rentabilität.

5. Sicherheit

Der Begriff Sicherheit beinhaltet den Schutz vor externen Angriffen (Security) und den Schutz vor Gefahren, die vom KI-System selbst ausgehen (Safety)⁴⁷. Die IT-Sicherheit (Security) umfasst die Bereiche Vertraulichkeit, Integrität und Verfügbarkeit. Wenn es aufgrund eines äusseren Angriffs zu einer Veränderung oder einer Beeinträchtigung der Funktionalität des KI-Systems kommt – beispielsweise wenn unautorisierte oder unbeabsichtigte Änderungen vorgenommen werden – betrifft dies die Integrität des KI-Systems. Beispiele sind «feindliche Attacken» (adversarial attacks) und Data Poisoning⁴⁶.

Bei einer «feindlichen Attacke» (adversarial attack) bewirkt ein Angreifer (Hacker) beim KI-System – z.B. durch manipulierte Daten – eine bestimmte Vorhersage. Hinter Data Poisoning steht die Absicht, Trainingsdaten gezielt zu manipulieren, um Funktionsänderungen herbeizuführen. Wenn das KI-System aufgrund eines äusseren Angriffs nicht mehr oder nur noch unzureichend reagiert, ist dessen Verfügbarkeit betroffen⁴⁶.

Der Bereich Safety umfasst Risiken, bei denen aufgrund einer Fehlfunktion oder des Ausfalls eines KI-Systems materielle Schäden entstehen oder Personen zu Schaden kommen können. Eine mögliche Ursache neben anderen ist mangelhaftes Design der Einbettung des KI-Systems in seine Umgebung⁴⁶.

6. Datenschutz

Hier geht es um den Schutz sensibler Daten bei der Entwicklung und während des Betriebes des KI-Systems. In die Kategorie der sensiblen Daten gehören in diesem Zusammenhang Personendaten sowie weitere, nicht personenbezogene Daten (z.B. Geschäftsgeheimnisse)⁴⁶. Aus datenschutzrechtlicher Sicht stehen bei der Anwendung von KI insbesondere die Fragen der Verantwortung, der Transparenz und des Einbezugs der betroffenen Person im Vordergrund. Es ist zu klären, ob der Einsatz von KI in Eigenverantwortung oder im Rahmen einer Auftragsdatenbearbeitung (z.B. Cloud-Lösung) erfolgt (Art. 9 neues Datenschutzgesetz DSG).

Schutz von KI-Systemen vor Sicherheitsrisiken zwingend

Für den Arzt bleibt die datenschutzrechtliche Verantwortung dieselbe, doch die Auftragsdatenbearbeitung bei KI erfordert spezifische vertragliche Regelungen (z.B. Verwendung der Daten durch Auftragnehmer für Entwicklung von KI-Systemen) (Art. 9 Abs. 1 lit.A neues DSG). Zu beachten gilt, dass auch ein KI-System durch geeignete technische und organisatorische Massnahmen gegen Sicherheitsrisiken angemessen zu schützen ist (Art. 8 neues DSG). Zudem ist vorgängig grundsätzlich eine Risikoabwägung (Datenschutz-Folgenabschätzung) durchzuführen, wenn eine Bearbeitung ein hohes Risiko für die Persönlichkeit oder die Grundrechte der betroffenen Person mit sich bringen kann (Art. 22 Abs. 1 neues DSG).

Vom KI-System abgefragte, verarbeitete oder gespeicherte Daten müssen wirksam geschützt werden.

Risikofaktoren sind «Data Poisoning» oder «Speech Synthesis». Dabei werden Daten manipuliert oder sprachliche Eingaben imitiert und somit verändert (siehe Dimension Sicherheit). Dies hat zur Folge, dass es zu einer Veränderung der Vorhersagen der KI-Systeme kommen kann. Gefahr droht auch von automatisiertem Hacking, bei dem geschützte sensible Daten (z.B. medizinische Patientendaten) an die Öffentlichkeit gelangen können⁷⁵. Diesbezüglich zu bedenken ist, dass Personen durch spezielle KI-Verfahren oder durch Verknüpfung (data linkage) der vom KI-System generierten Daten mit weiteren Daten unter Umständen re-identifizierbar sind, was zu einer Verletzung der Persönlichkeitsrechte der Betroffenen führen kann. Es ist somit notwendig, die vom KI-System abgefragten, verarbeiteten oder gespeicherten Daten während der Entwicklung und während des Betriebes wirksam zu schützen⁴⁶.

Strengere Vorgaben im neuen Datenschutzgesetz DSG

Bei der Anwendung von KI muss bei der Verwendung von Personendaten im Kontext einer Behandlung ein Rechtfertigungsgrund (i.e. Einwilligung) vorliegen (Art. 31 neues DSG), oder es muss mit

anonymisierten Daten gearbeitet werden. Generell muss auch bei der Verwendung von KI die Datenbearbeitung für die betroffenen Personen erkennbar sein (Art. 6 Abs. 3 neues DSGVO; Art. 19 neues DSGVO). Für die Verwendung von KI kann eine ausdrückliche Einwilligung eingeholt werden. Liegt keine solche vor, ist die betroffene Person über eine automatisierte Einzelentscheidung zu informieren, sofern diese mit einer Rechtsfolge verbunden oder die Person erheblich beeinträchtigt ist und der Entscheid ausschliesslich auf KI beruht (Art. 21 neues DSGVO). Die betroffene Person kann auch Auskunft über die Logik der automatisierten Entscheidung verlangen (Art. 25 Abs. 2 lit. F neues DSGVO).

Neue technische Ansätze für die datenschutzkonforme Bearbeitung grosser Datensätze

Das Zusammenführen von Personendaten zu grossen Datensätzen, sogenannte «Big Data»-Datensätze, stellt aus datenschutzrechtlicher Sicht eine Herausforderung dar, die dem Teilen von Daten Grenzen setzt. KI-Systeme benötigen allerdings für das Trainieren ihrer Modelle grosse für die Patientenpopulation relevante Datensätze.

Als Lösung bieten sich neue technische Ansätze an, die es KI-Systemen erlauben, auf verteilte Datensätze zuzugreifen, ohne dass diese physisch zusammengelegt werden müssen. Ein Beispiel für eine solche Technologie ist das föderierte Lernen⁷⁶: In einem ersten Schritt werden die KI-Modelle auf verschiedenen lokalen Spitaldaten trainiert. In einem zweiten Schritt erfolgt durch physisches Teilen der Modellparameter die zentrale Zusammenführung der lokalen Modellparameter in einem globalen Modell. Es hat sich gezeigt, dass so trainierte globale Modelle ähnliche Leistungen erbringen wie Modelle, die direkten Zugriff zu allen Daten (d. h. zu einem «Big Data» Datensatz) haben.

Eine Weiterentwicklung dieser Technologie ist das sogenannte swarm learning. Dieses verwendet ebenfalls verteilte Datensätze für das Trainieren der KI-Modelle, kommt aber ohne ein zentral geführtes Modell aus⁷⁷.

Bei beiden Technologien greifen die Datenschutzbestimmungen nur bedingt, da die Daten die Spitäler nicht verlassen und lediglich abstrakte und anonymisierte Informationen (Modellparameter) geteilt werden.

Aufklärung, Patienteneinwilligung und Dokumentation

- **Aufklärung und Einwilligung** Es ist davon auszugehen, dass beim Einsatz von KI in Zukunft erhöhte Anforderungen an die Aufklärungspflichten hinsichtlich «neuer unbekannter Risiken» gestellt werden. Der Patient soll via Aufklärung sein Selbstbestimmungsrecht wahrnehmen und den medizinischen Eingriff einschliesslich des Einsatzes von KI verstehen können. Nur so ist er in der Lage, rechtsgültig die Einwilligung dazu erteilen^{d,e} (vgl. auch Art. 10 der Standesordnung der FMH⁸³).
- **Dokumentation** Die Behandlungsschritte inklusive Einsatz von KI sowie die Aufklärungsgespräche sind zu dokumentieren und in die Krankengeschichte (KG) einzutragen. Die ärztliche Dokumentation dient einerseits der Therapiesicherung und andererseits der Beweissicherung. Die KG muss so abgefasst sein, dass der medizinische Behandlungsablauf vollständig und nachvollziehbar dokumentiert ist. Aufzeichnungspflichtig sind somit die für die ärztliche Behandlung wesentlichen medizinischen Fakten^f (vgl. auch Art. 12 der Standesordnung der FMH⁸³).

^d BGE 133 III 121

^e BGE 117 Ib 197

^f BGE 141 III 363

KI-Modelle im Kontext der Frage des geistigen Eigentums

Auch nicht-personenbezogene Daten können schützenswert sein, also alle Arten von Daten an denen Dritte Rechte haben wie beispielsweise lizenzgebundene Daten. Ebenfalls kann das KI-Modell innerhalb eines KI-Systems schützenswert sein, wenn es geschäftsrelevant ist⁴⁶. Geistiges Eigentum (Intellectual Property, IP) sichert menschliche Innovationen und Erfindungen. Da KI-Systeme ebenfalls das Potenzial haben, Neues zu erschaffen, stellen sich in Bezug auf das geistige Eigentum folgende Fragen: Wie sollen menschlich erschaffene Innovationen und Erfindungen gegenüber den durch KI-Systeme erzeugten bewertet werden? Ist es nötig, die bestehenden IP-Richtlinien für KI-Systeme anzupassen, um beispielsweise die von KI geschaffenen Erfindungen, die KI selbst und die für KI verwendeten Daten zu schützen⁷⁸?

Aktuelle Rechtslage beim Urheberrecht

Gemäss dem im IP-Bereich allgemein geltenden Schöpferprinzip stehen Immaterialgüterrechte ohne anderweitige gesetzliche Grundlage dem Schöpfer des betreffenden Immaterialguts zu. Als Schöpfer kommen ausschliesslich natürliche Personen in Betracht. Folglich werden mit KI-Systemen geschaffene Immaterialgüter herkömmlicherweise natürlichen Personen zugeordnet. Für das Urheberrecht ist das Vorliegen einer geistigen Schöpfung, d. h., eines Ausdrucks des menschlichen Geistes, sogar eine explizite Schutzvoraussetzung. Der Fokus dürfte auf diejenigen natürlichen Personen gerichtet sein, die das KI-System programmiert und insbesondere mit Trainingsdaten und anderen Inputs und Vorgaben parametrisiert haben. Lässt sich aufgrund der weitgehenden Autonomie des KI-Systems hingegen keine natürliche Person als Schöpfer identifizieren, ist die Schöpfung gemäss derzeitiger Rechtslage in aller Regel nicht schutzfähig⁷⁹. Aufgrund der bestehenden rechtlichen Unsicherheiten, empfiehlt es sich, auf klare vertragliche Vereinbarungen unter den Beteiligten über die Zuordnung von allenfalls entstehenden Immaterialgüterrechten zu achten.

Zwischenfazit

Welcher menschliche Akteur muss dafür Sorge tragen, dass ein KI- System vertrauenswürdig ist und bleibt?

Beauchamp und Childress haben vier medizinethische Grundprinzipien definiert: Respekt vor der Autonomie des Patienten, Schadensvermeidung, Fürsorge und Gerechtigkeit⁸⁰. Ethik-Leitlinien für vertrauenswürdige KI wurden auch von der EU entwickelt. Sie stützen sich auf vier Prinzipien des ethischen Handelns: Achtung der menschlichen Autonomie, Schadensverhütung, Fairness und Erklärbarkeit⁴⁵. Die Prinzipien und Leitlinien verfolgen das Ziel, dass ein KI-System für vergleichbare Patientengruppen identische Vorhersagen (Generalisierbarkeit) mit einer vergleichbaren und robusten Performanz (Treffericherheit/Accuracy, Sensitivität, Spezifität) trifft und medizinische Interventionen fair zuteilt. Dies unabhängig von sozioökonomischen Faktoren⁴⁹. Solange KI-Systeme – beispielsweise in Form einer elektronischen Person (E-Person^{81,82}) – keine Rechtspersönlichkeit besitzen, stehen diejenigen in der Verantwortung, die es konzipieren und professionell nutzen, also auch Gesundheitsfachpersonen. Sie müssen darauf achten, dass die KI-Systeme den Prinzipien menschlichen Handelns nicht zuwiderlaufen (vgl. auch Art. 2–4 der Standesordnung der FMH⁸³).

Wie wird künstliche Intelligenz in der Medizin reguliert?

Bis heute gibt es keine Gesetze oder international einheitliche Richtlinien für die Regulierung des Einsatzes von KI in der Medizin. Verschiedene Behörden und Organisationen sind jedoch im Begriff, entsprechende Regularien und Vorschriften zu entwickeln oder bestehende Gesetze und Normen anzupassen.

Die FDA hat 2019 ein Diskussionspapier veröffentlicht, in dem sie regulatorische Rahmenbedingungen für KI/ML-basierte Software in Medizinprodukten vorstellt⁸⁴. Basierend auf den Rückmeldungen wurde ein Aktionsplan für diese Anwendungen formuliert⁸⁵. Thematisiert wird beispielsweise, wie mit KI-Systemen umgegangen werden könnte, wenn diese kontinuierlich weiterlernen. Alle bisher von der FDA zugelassenen KI-basierten Medizinprodukte sind sogenannte «eingefrorene» Algorithmen (locked algorithms), die nicht kontinuierlich weiterlernen⁸⁴. Die von der FDA zugelassenen KI/ML-basierten Technologien sind in einer Open-Access-Datenbank zusammengefasst⁵⁵. 2021 hat die FDA 10 Leitprinzipien für die Entwicklung von ML-basierten Medizinprodukten publiziert⁸⁶.

Die EU hat im April 2021 einen Vorschlag für die Festlegung harmonisierter Vorschriften für KI (Gesetz über KI) in die Vernehmlassung gegeben⁸⁷. Im November 2021 wurde der Vorschlag von der Kommission mit Änderungen angenommen⁸⁸.

Die gesetzlichen Rahmenbedingungen und Regularien für den Einsatz von KI sind erst im Entstehen begriffen.

Das Dokument definiert KI samt der möglichen zugrunde liegenden Methodik (Art. 3 und Anhang I⁸⁷). Die EU klassifiziert die KI-Systeme nach ihrem Risiko (unannehmbares, hohes, geringes, minimales) und definiert je nach Risikogruppe unterschiedliche Anforderungen. Der Hochrisikogruppe zugeordnet sind medizinische KI-Systeme, die Medizinprodukte der Klassen Im, Is, Ila, Ilb, II nach MDR⁸⁹ darstellen (einzig die Klasse I fällt nicht darunter), diejenigen, die unter die In Vitro Diagnostika Verordnung (IVDR) fallen (Art. 6, Anhang II⁸⁷), sowie alle KI-Systeme für die Priorisierung der Not- und Rettungsdienste (Anhang III⁸⁷).

Ein neuer, von der EU-Kommission eingefügter Art. 52a und der damit verbundene neue Erwägungsgrund in Art. 70a sollen klarstellen, dass KI-Systeme für allgemeine Zwecke nicht als bestimmungsgemäss im Sinne dieser Verordnung angesehen werden sollten⁸⁸. Der EU-Vorschlag sieht auch einen Europäischen Ausschuss für KI vor. Dieser soll sich aus Vertretern der nationalen für KI zuständigen Aufsichtsbehörden, dem EU Datenschutzbeauftragten und der Kommission zusammensetzen und die reibungslose, wirksame und einheitliche Umsetzung der KI-Verordnung garantieren⁹⁰. 2021 hat die EU zudem eine gesetzesvergleichende Studie zur zivilrechtlichen Haftung von KI veröffentlicht⁹¹.

Daneben gibt eine Reihe von weiteren internationalen Empfehlungen und Normen zu verschiedenen Aspekten des Einsatzes von KI-Systemen sowie für speziell auf Machine Learning zugeschnittene Systeme⁹². Ausserdem existiert ein von einem interdisziplinären Team entwickelter Leitfaden für Hersteller von KI-Systemen und Auditoren, also jene Stellen, die KI-Systeme als Medizinprodukt bewerten müssen. Dieser Leitfaden stellt eine Handlungsanleitung dar, um einerseits KI-Systeme regelkonform zu entwickeln und als Medizinprodukt im Europäischen Raum zuzulassen, und andererseits mit Hilfe von Checklisten deren Regel-Konformität zu überprüfen⁹³.

Geltende Rechtsordnung auch für den Einsatz von KI verbindlich

Um einige Herausforderungen, die der Einsatz von KI für die globale Arzneimittelregulierung mit sich bringt zu verdeutlichen, hat die Internationale Koalition der Arzneimittelbehörden (ICMRA) zwei hypothetische KI-Fallstudien entwickelt. Anhand dieser Studien prüfte sie, ob und inwiefern die bestehenden Regularien in Bezug auf KI angepasst werden müssten⁹⁴. Auch für die Schweiz gilt: Der Einsatz von KI hat sich in allen Aspekten an die geltende Rechtsordnung zu halten. Wo diese den technologischen Entwicklungen nicht gerecht wird – so z. B. beim Einsatz von KI «Out of the Loop», respektive bei KI, deren Ergebnisse keinem menschlichen Urheber mehr zugerechnet werden können – ist sie entsprechend anzupassen und/oder zu ergänzen⁸⁵. Aktuell wird national diskutiert, welche rechtlichen Ansätze zur Regulierung von KI-Systemen in der Schweiz verfolgt werden sollten^{95,96,97}. Wünschenswert zur Klärung dieser Frage ist ein gesellschaftlicher Diskurs zur Ziehung möglicher und gegebenenfalls notwendiger Grenzen zwischen dem technisch Machbaren und ethischen sowie finanziellen Gesichtspunkten.

Welche Forderungen stellt die FMH an die künstliche Intelligenz in der Medizin?

Der Einzug von künstlicher Intelligenz in der Medizin erschliesst für die Tätigkeit von Ärztinnen und Ärzten interessante Perspektiven. Allerdings wirft die Anwendung von KI im medizinischen Alltag auch viele Fragen auf. Deshalb hat die FMH einen Forderungskatalog erstellt, der zu einem sinnvollen und zweckmässigen Einsatz von KI beitragen soll.

Künstliche Intelligenz dürfte das Denken und Handeln der Ärzteschaft in den kommenden Jahren tiefgreifend verändern. Mit Hilfe von KI soll es gelingen, alle Prozesse entlang der Patient Journey sinnvoll und zweckmässig weiterzuentwickeln und zu verbessern. Um die damit verbundenen Herausforderungen besser zu meistern, hat die FMH aus Sicht der Ärzteschaft **zehn Forderungen** formuliert.

Die **zehn Forderungen** beziehen sich auf diejenigen Systeme, die zu diagnostischen respektive therapeutischen Zwecken herangezogen werden und damit das Potenzial haben, ärztliche Kerndienstleistungen zu ersetzen. Der Fokus liegt auf KI-Systemen, die gemäss Medizinprodukteverordnung (Medical Device Regulation, MDR) nach Klasse IIA und höher eingestuft werden.

Einbezogen wurden die folgenden bestehenden oder vorgeschlagenen Gesetze, Regularien und Leitfäden: MDR⁸⁹, MepV⁹⁸, IVDR⁹⁹, Gesetz über KI⁸⁷, Leitfaden zur KI bei Medizinprodukten⁹³ sowie die Guiding Principles for Good Machine Learning Practice for Medical Device Development⁸⁶.

- 1. Das KI-System soll die menschliche Intelligenz und die Beziehung zwischen Ärztin oder Arzt und Patientin oder Patient stärken und nicht ersetzen¹⁰⁰.**
- 2. Das KI-System muss sich an den Prinzipien der evidenzbasierten Medizin (EbM)¹⁰¹ orientieren.**
- 3. Die Leistungen eines KI-Systems müssen wirkungsvoll und nutzbringend sein. Sie sollen administrative Prozesse vereinfachen und Ärztinnen und Ärzte davon entlasten.**
- 4. Das KI-System muss regelmässig überprüft werden, und es sind unverzüglich Korrekturen vorzunehmen, wenn sich solche aufgrund der Überprüfung als geboten erweisen.**

Das KI-System muss im Sinne einer Post-Market-Surveillance regelmässig überprüft werden. Dabei geht es insbesondere darum, neue Risiken zu identifizieren oder die Leistungsfähigkeit sowie das Nutzen-Risiko-Verhältnis in der realen klinischen Umgebung zu überprüfen und falls notwendig kontinuierlich zu verbessern. Z. B. kann ein KI-System belastbare Vorhersagen nur für eine Patientenpopulation mit einer Anzahl von Eigenschaften treffen, für die es trainiert, getestet und validiert wurde.
- 5. Ärztinnen und Ärzte bleiben in Abstimmung mit ihren Patientinnen und Patienten Entscheidungsträger in Bezug auf den Einsatz und den Umgang mit dem KI-System.**

Solange KI-Systeme – beispielsweise in Form von elektronischen Personen (E-Person^{81, 82}) – keine Rechtspersönlichkeit besitzen, stehen diejenigen in der Verantwortung, die es konzipieren und professionell nutzen, also auch Gesundheitsfachpersonen.

Selbst bei «Out of the Loop»-Systemen, bei denen menschliches Eingreifen nicht mehr erforderlich ist, kann der Nutzer noch entscheiden, ob er das KI-System einsetzen möchte oder nicht, auch wenn er die Vorhersagen und Entscheidungen dieses autonomen KI-Systems selbst nicht mehr beeinflussen kann.

6. Dem KI-System muss eine «Gebrauchsanleitung» für Ärztinnen und Ärzte beiliegen.

Die «Gebrauchsanleitung» muss alle erforderlichen Informationen zur einwandfreien Verwendung des KI-Systems beinhalten. Sie muss Ärztinnen und Ärzte in die Lage versetzen, den Betrieb des KI-Systems zu verstehen, zu überwachen und bei Bedarf auch zu übersteuern. Der Hersteller muss sicherstellen, dass Ärztinnen und Ärzten bei der Verwendung des KI-Systems ihren Informations- und Sorgfaltspflichten nachkommen können.

Die «Gebrauchsanleitung» muss klar aufzeigen:

... welche Rolle der Ärztin oder dem Arzt beim Betrieb des Systems zukommt.

Definiert werden müssen z. B. die Nutzerkompetenzen bei «In the Loop»- oder «On the Loop»-Systemen.

... für welchen Zweck bzw. für welches Krankheitsbild das KI-System eingesetzt werden darf und für welche Anwendungsfälle nicht.

... wie das KI-System zu verwenden ist.

Speziell muss aufgezeigt werden, welche Daten wie eingegeben werden müssen, wie Fehlfunktionen und besondere Vorfälle erkannt werden können, woran erkannt werden kann, dass das System einwandfrei funktioniert, und wie das System übersteuert werden kann.

... auf welcher Methodik das KI-System basiert.

Zu beachten und umzusetzen sind die Grundprinzipien und Goldstandards der Datenkompetenz.

Dazu gehören beispielsweise ein sinnvolles transparent deklariertes und nachvollziehbares Datenkonzept oder eine dem Konzept angepasste Datenerhebung und Datenverarbeitung.

... wie das KI-System trainiert, getestet und validiert wurde.

Anzugeben sind insbesondere die jeweils verwendeten unabhängigen Datensätze und deren Charakteristika wie Grösse, Verteilung Geschlecht, Alter usw. Des Weiteren die Eignung der Datensätze aus statistischer und medizinischer/klinischer Perspektive inklusive Beschreibung der Tests auf mögliche Verzerrungen (Bias) und Fehler. Aufzuführen sind ausserdem die Kennzahlen zur Genauigkeit der Leistung und der klinischen Wirksamkeit des KI-Systems, auch im Vergleich zu bestmöglichen alternativen Methoden und KI-Modellen.

... wie das KI-System den Datenschutz und die Sicherheit gewährleistet.

Insbesondere ist darzulegen, wo und in welcher Form die Personen- und Metadaten gespeichert werden, wie diese vor Sicherheitsrisiken angemessen geschützt werden und wie die Sicherheit des KI-Systems gewährleistet wird.

... an welche Kontaktstelle sich Nutzerinnen und Nutzer wenden können.

Es ist sicherzustellen, dass die Ärztin oder der Arzt im Falle von Fragen oder Problemen kompetente Hilfestellung erhält.

7. Ärztinnen und Ärzten ist für den Einsatz und Umgang mit KI-Systemen ein angemessenes Aus-, Weiter- und Fortbildungsangebot zur Verfügung zu stellen.

Das medizinische Ausbildungssystem muss angehenden und praktizierenden Ärztinnen und Ärzten das Werkzeug in die Hand geben, um im ärztlichen Alltag mit KI und modernen Techniken sinnstiftend und kompetent umzugehen. Dies schliesst die Auseinandersetzung mit rechtlichen und ethischen¹⁰² Rahmenbedingungen ebenso mit ein wie die Befähigung, kompetent mit Daten umzugehen und die verschiedenen Aspekte der datengetriebenen Medizin zu verstehen und zu begleiten^{103,104,105,106}. Damit ein informierter Umgang mit KI möglich wird, ist es beispielsweise unabdingbar, dass grundlegende Begriffe der Datenanalyse auch in Bezug auf das Machine Learning verinnerlicht werden.

8. Bei der Entwicklung von KI-Systemen müssen Ärztinnen und Ärzte sowie Patientinnen und Patienten in die Definition der Nutzungsanforderungen einbezogen werden.

Damit die Wirkungsziele der Nutzerinnen und Nutzer erreicht werden können, ist insbesondere zu spezifizieren, welche medizinischen Fragestellungen mit dem KI-System gelöst werden sollen, worin die medizinischen Probleme liegen und in welchem Kontext das KI-System verwendet werden soll.

9. Ärztinnen und Ärzte müssen über den Einsatz von KI-Systemen, die ihre Arbeit indirekt beeinflussen, informiert werden.

Denkbar sind KI-Systeme, die im Hintergrund die Qualität der ärztlichen Arbeit überwachen.

10. Es muss eine nationale Dateninfrastruktur geschaffen werden, die es ermöglicht die KI-Systeme mit grossen qualitativ hochwertigen annotierten medizinischen Datensätzen zu trainieren, zu testen und die Reproduzierbarkeit der Ergebnisse durch eine robuste Validierung sicherzustellen.

KI-Systeme müssen auf die lokale Datensituation angepasst werden (s. Forderung 4). Dazu braucht es Datensätze, die die jeweilige zu testende lokale Patientenpopulation repräsentieren und die unter Beachtung des Datenschutzes und der Datensicherheit erhoben, verarbeitet und gespeichert werden.

Wichtig ist in diesem Zusammenhang ein zertifiziertes Qualitätsmanagement und ein zertifiziertes Annotieren der verwendeten und gesammelten Daten.

Darüber hinaus braucht die Schweiz Zugang zu internationalen Datenplattformen^h. Die Schweizerische Ärzteschaft kann sich hier auch auf neue Lösungen, wie etwa die des föderierten Lernens, abstützen, die ein dezentrales Lernen von Daten ermöglichen.

Die FMH steht mit ihren Forderungen nicht alleine da. So haben die europäische Ärztesgesellschaft (CPME, Standing Committee of European Doctors) und die amerikanische Ärztesgesellschaft (AMA, American Medical Association) minimale Forderungen an KI-Systeme definiert^{108, 109}.

Forderungen an KI waren in der Schweiz auch Thema von Workshops. Beispielsweise unter dem Motto «Regulierung von KI getriebenen Algorithmen zur Entscheidungsfindung» (IRGC, International Risk Governance Center) oder zur Frage «KI in der Medizin – kann man ihr vertrauen» (SAMW, Schweizerische Akademie der medizinischen Wissenschaften; SATW, Schweizerische Akademie der technischen Wissenschaften)^{110, 111}.

^h Als Beispiel sei das Vereinigte Königreich angeführt, wo im CPRD Netzwerk Daten von Hausärzten gesammelt werden.¹⁰⁷

Wie wird sich das Berufsbild Ärztin/Arzt in Zukunft wandeln?

Der Einzug der künstlichen Intelligenz in der Medizin bewirkt mittelfristig einen tiefgreifenden Wandel des Berufsbildes der Ärztinnen und Ärzte. Allerdings wird künstliche Intelligenz ihre Rollen kaum je vollumfänglich übernehmen.

KI hat das Potenzial, die Qualität der medizinischen Versorgung zu verbessern. Sie kann sowohl menschliche Fehler reduzieren als auch die Belastung der Ärztinnen und Ärzte durch Routineaufgaben minimieren, beispielsweise im administrativen Bereich. Wenn es also gelingt, KI für administrative und medizinische Routineaufgaben erfolgreich in den klinischen Alltag zu integrieren, könnte dies den Ärztinnen und Ärzten mehr Zeit für anspruchsvolle Aufgaben und für den persönlichen Kontakt mit ihren Patientinnen und Patienten verschaffen³⁹. Damit würde sich das Berufsbild Ärztin/Arzt tatsächlich stark verändern. Die Frage ist, in welchem Ausmass.

Zunächst müssen die Rollen und damit verbunden die Fähigkeiten betrachtet werden, die eine Ärztin oder ein Arzt tagtäglich anwenden respektive in der Ausbildung erwerben muss, um diesen Beruf bestmöglich auszuüben.

Definition der verschiedenen ärztlichen Rollen durch CanMEDS

Eine Definition dieser Rollen hat das CanMEDS Rahmenwerk (Canadian Medical Education Directives for Specialists (CanMEDS) Physician Competency Framework¹¹²) geliefert. Es beschreibt, welche Rollen eine Ärztin oder ein Arzt einnehmen sollte und welche Fähigkeiten dafür notwendig sind. Dieses Rahmenwerk wurde in den Schweizer Lernzielkatalog für Medizinstudentinnen und -studenten aufgenommen.

Der Ärztin oder dem Arzt werden nach CanMEDS sieben Rollen zugeschrieben. Sie oder er agiert als Medizinexperte, Kommunikator, Kollaborator, Führungsperson, Gesundheitsfürsprecher, Gelehrter und Fachkraft. Dabei schliesst die Rolle des Medizinexperten die übrigen sechs Rollen mit ein.

Keine dieser Rollen des ärztlichen Berufsbildes ist ausschliesslich technisch definiert. Es ist deshalb unwahrscheinlich, dass intelligente Maschinen diese Rollen vollständig übernehmen können¹¹³. Allerdings verfügen intelligente Maschinen über fast unbegrenzte Rechenleistung und Speicherkapazitäten. Diese befähigen sie dazu, extrem grosse multimodale Datensätze extrem schnell zu verarbeiten und komplexe Algorithmen anzuwenden. Intelligente Maschinen sind somit in der Lage, komplexe Hypothesen zu testen und zwar zu jeder Tages- und Nachtzeit. Im Vergleich zu Menschen können sie Informationen systematischer verarbeiten, Entscheidungen konsequenter treffen und schneller auf Veränderungen reagieren².

Einiges können intelligente Maschinen besser als der Mensch. Anderes nicht. Technik kann zwischenmenschliche Beziehungen nicht ersetzen.

Da eine vollständige Übernahme der ärztlichen Rollen durch intelligente Maschinen kaum realistisch ist, stellt sich die Frage, welche Kompetenzen daraus im medizinischen Alltag auf KI-Systeme übertragen werden könnten, und wie sich dies auf die Rollen des ärztlichen Berufsbildes auswirken würde.

Definition der CanMEDS Rollen¹¹²

Begriff	Definition
Medizinexperte <i>medical expert</i>	Die Rolle des Medizinexperten ist im Ärzteberuf zentral und bezieht sich auf den klinischen Tätigkeitsbereich. Als Medizinexperten vereinen Ärztinnen und Ärzte alle weiteren CanMEDS Rollen, indem sie medizinisches Wissen, klinische Fähigkeiten und berufliche Werte bei der Bereitstellung einer hochwertigen und sicheren patientenorientierten Versorgung anwenden.
Kommunikator <i>communicator</i>	Als Kommunikatoren bauen Ärztinnen und Ärzte Beziehungen zu ihren Patientinnen und Patienten und deren engem Umfeld (z. B. Familien) auf. Diese Beziehung erleichtert es, wichtige Informationen für eine effektive Gesundheitsversorgung zusammenzutragen und auszutauschen.
Kollaborator <i>collaborator</i>	Als Kollaborateure arbeiten Ärztinnen und Ärzte effektiv mit anderen Gesundheitsfachkräften zusammen, um eine sichere, hochwertige und patientenorientierte Versorgung zu gewährleisten.
Führungsperson <i>leader</i>	Als Führungsperson engagieren sich Ärztinnen und Ärzte gemeinsam mit anderen für die Vision eines qualitativ hochwertigen Gesundheitssystems. Sie übernehmen durch ihre Tätigkeit als Ärzte, Verwalter, Gelehrte oder Lehrer Verantwortung für die Bereitstellung einer hervorragenden Patientenversorgung.
Gesundheitsfürsprecher <i>health advocate</i>	Als Gesundheitsfürsprecher bringen Ärztinnen und Ärzte ihr Fachwissen und ihren Einfluss ein, wenn sie mit anderen Stakeholdern oder Patientengruppen kooperieren, um das Gesundheitswesen zu verbessern. Sie arbeiten mit Patientinnen und Patienten zusammen, um deren Bedürfnisse zu ermitteln und zu verstehen. Sie sprechen bei Bedarf für andere und unterstützen die Mobilisierung von Ressourcen, um Veränderungen zu bewirken.
Gelehrter <i>scholar</i>	Als Gelehrte zeigen Ärztinnen und Ärzte ein lebenslanges Engagement für hervorragende Leistungen in der Praxis durch kontinuierliches Lernen. Auch indem sie andere unterrichten, Erkenntnisse bewerten und zur Lehre beitragen.
Fachkraft <i>professional</i>	Als Fachkraft sind Ärztinnen und Ärzte der Gesundheit und dem Wohlergehen der einzelnen Patientinnen und Patienten und der Gesellschaft verpflichtet. Grundpfeiler sind ethische Verhaltensweisen, hohe persönliche Verhaltensstandards, Verantwortlichkeit gegenüber dem Berufsstand und der Gesellschaft, ärztlich gelenkte Regulierung sowie die Aufrechterhaltung der persönlichen Gesundheit.

Die neue Rolle der Ärztin oder des Arztes ...

... als Medizinexperte:

Die Stärke von KI-Systemen liegt wie erwähnt in ihrer Fähigkeit, sehr grosse Datenmengen mit hoher Komplexität verarbeiten zu können. Einerseits rufen sie Expertenwissen ab, andererseits generieren sie Expertenwissen aus eigener Kraft. Bei beiden Punkten handelt es sich um Teilaspekte der Rolle des Medizinexperten. Doch das Wesen des Medizinexperten besteht darin, nicht nur Wissen zu verarbeiten, wie das ein KI-System zum Teil auch kann,

sondern ebenso darin, alle anderen Rollen des Arztberufes mit einzubeziehen (Kommunikator, Kollaborator, Führungsperson, Gesundheitsfürsprecher, Gelehrter, Fachkraft). Dazu wird KI wahrscheinlich nur teilweise fähig sein.

... als Kommunikator und Kollaborator:

Die Rolle des Kommunikators wird an Bedeutung gewinnen. Wie positiv oder negativ Patientinnen und Patienten den Einsatz von KI zur Unterstützung ihrer Gesundheit beurteilen, dürfte höchst unterschiedlich sein. Die Ärztin oder der Arzt ist somit gefordert, auf individuelle Fragen – beispielsweise zum Datenschutz oder zur Nachvollziehbarkeit – einzugehen und neutral sowie ergebnisoffen zu

informieren¹¹³ (siehe auch Kapitel Herausforderungen: Stichwort Einwilligung der Patientinnen und Patienten).

Da durch die Digitalisierung im Gesundheitsbereich medizinisches Wissen zugänglicher wird, bleibt Expertenwissen nicht mehr ausschliesslich Ärztinnen und Ärzten vorbehalten, sondern ist auch für andere Gesundheitsfachpersonen wie Psychologen, Krankenschwestern usw. sowie für die Patientinnen und Patienten zugänglich. KI-Systeme unterstützen zunehmend die Verfügbarkeit des medizinischen Wissens für diese Personengruppen. Beispiele sind das weiter oben erwähnte KI-gesteuerte Ultraschallsystem oder die Smartwatch zur Erkennung von Arrhythmien. Nicht-ärztliche Gesundheitsfachpersonen wie auch Patientinnen und Patienten können sich somit eigenständige Meinungen über einen Gesundheitszustand bilden. Dieser Aspekt wird weitreichende Folgen für die ärztliche Rolle des Kollaborators und Kommunikators haben¹¹³.

... als Führungsperson und Gesundheitsfürsprecher:

Als Führungsperson sollte die Ärztin oder der Arzt mit Patientinnen und Patienten zusammenarbeiten. Dies mit dem Ziel, einen transparenten Einsatz von KI in der Medizin zum Wohle der Patientinnen und Patienten zu fördern. Als Gesundheitsfürsprecher können sie zur Klärung der Frage beitragen, ob medizinische KI-Anwendungen soziale Ungerechtigkeiten verringern oder verstärken. Gegebenenfalls ist es angezeigt, die Interessen benachteiligter Patientengruppen wahrzunehmen.

... als Gelehrter:

Voraussetzung für die Wahrnehmung all dieser veränderten ärztlichen Rollen, d. h., für den kompetenten Umgang mit KI ist eine hohe digitale Kompetenz der Ärztinnen und Ärzte. Sie müssen die Fähigkeit und den Willen aufbringen, kontinuierlich zu Lernen: über Herkunft, Aufbereitung, Integration und Governance von Gesundheitsdaten sowie über Methodik von KI, Ethik von KI und deren kritische Bewertung und Interpretation^{113,4}. Diese Lernbereitschaft ist eine Kernaufgabe im Rahmen der veränderten ärztlichen Rolle des Gelehrten.

... als Fachkraft:

Als Fachkraft sollte die Ärztin oder der Arzt akzeptieren, dass KI sie in gewissen Teilbereichen zukünftig unterstützen wird. Dies erlaubt es ihnen, sich auf ihre Kernaufgaben zu konzentrieren beziehungsweise diejenigen Kompetenzen weiterzuentwickeln, die ein KI-System eben nicht übernehmen kann, zumindest nicht in naher Zukunft¹¹³.

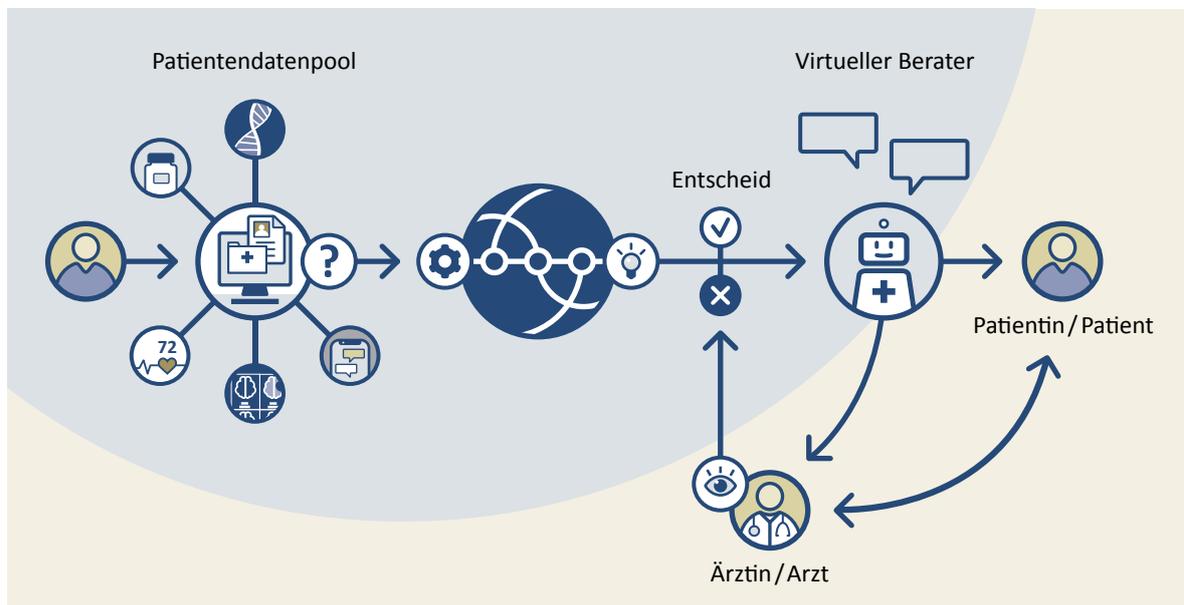
Ärztinnen und Ärzte als Informationsverarbeitende, Dolmetscher und Berater

Der Einzug von KI in der Medizin bedeutet für das ärztliche Berufsbild, dass die medizinische Gesundheitsversorgung immer mehr zu einem Zusammenspiel von Ärztinnen und Ärzten, weiteren Gesundheitsfachpersonen, KI-Systemen sowie Patientinnen und Patienten werden wird. Die Erbringung medizinischer Leistungen erfolgt in Zukunft zunehmend da, wo sich der Patient gerade befindet, also an vielen verschiedenen Orten. Beispielhaft für diese Entwicklung ist der Einsatz der Telemedizin während der Covid-Pandemie.

Die medizinische Technologie wird sich mit und in den Patientinnen und Patienten bewegen und kontinuierliche Datenströme liefern, die es einzuordnen gilt¹¹⁴. Dadurch werden medizinische Fachdisziplinen entstehen, die wiederum die Zusammensetzung des Fachpersonals im «Care-Team» verändern. Auch die kompetente, zweckmässige und effiziente Gestaltung der Interaktion zwischen Mensch und intelligenter Maschine wird in diesem Zusammenhang ihren Beitrag leisten. Ärztinnen und Ärzte werden sich in naher Zukunft eher auf jene Bereiche ihrer neuen Rollen konzentrieren dürfen, die KI-Systeme nicht übernehmen können. Sie werden die Funktion eines Dreh- und Angelpunkts für medizinische, technische, regulatorische und ethische Informationsverarbeitung wahrnehmen und gleichzeitig Dolmetscher und Patientenberater sein³⁹.

Postuliert wird auch, dass in den kommenden²⁰ Jahren virtuelle medizinische Berater ganzheitliche personalisierte Gesundheitsberatung und Krankheitsprävention ermöglichen. Sie greifen auf verschiedenste medizinische und persönliche Daten (z. B. Genomanalyse, Mikrobiomanalyse, Ernährung, physische Aktivität usw.) zu, erfassen diese und werten sie mit Hilfe von KI aus (Abb.7). Diese Vision soll durch die Kombination von Genomanalyse mit digitaler Medizin, KI und Robotik^{3,4} Realität werden.

Abb. 7 Der virtuelle medizinische Berater, der multimodale Daten erfasst und diese mit Hilfe von KI auswertet.



- ❓ Problem- oder Fragestellung für das KI-System
- ⚙️ Verarbeitung der Daten im KI-System
- 💡 Lösungsvorschlag des KI-Systems zur Problem- oder Fragestellung
- 👁️ Überprüfung des KI-Entscheids durch Ärztin/Arzt
- ✅❌ Genehmigung/Ablehnung des KI-Lösungsvorschlags durch KI-System

Neuausrichtung der ärztlichen Ausbildung

In jedem Fall muss das medizinische Ausbildungssystem angehenden und praktizierenden Ärztinnen und Ärzten das Werkzeug in die Hand geben, um mit KI und modernen Techniken sinnstiftend umzugehen. Dies schliesst die rechtlichen und ethischen¹⁰² Rahmenbedingungen mit ein. Derzeit werden Studierende auf die durch CanMEDs definierten Rollen vorbereitet. Diese widerspiegeln allerdings eine Medizin, die weitgehend noch ohne KI-Systeme betrieben wurde. Nun zeichnet sich ab, dass Ärztinnen und Ärzte in Zukunft auch auf eine weitere Rolle vorbereitet werden müssen: die des Datenanalysten. Diese beinhaltet die Befähigung, die verschiedenen Aspekte der Daten-getriebenen Medizin zu verstehen und zu begleiten. So gilt es z. B., grundlegende Begriffe der Datenanalyse auch in Bezug auf das Machine Learning zu verinnerlichen, damit ein informierter Umgang mit der KI möglich wird.

Bestrebungen für eine Reform des Medizinstudiums in diese Richtung sind im Gange. Bereits hat das Schweizer Bildungsnetzwerk Medizin diesbezügliche Lernziele für Medizinstudentinnen und -studenten definiert^{104,103}. Einige Schweizer Ausbildungsstätten bieten schon heute Aus-, Weiterbildungs- und Fortbildungsangebote für Medizinstudentinnen und -studenten sowie Ärztinnen und Ärzte an^{115,116}. Auch die FMH wird sich dem Thema «Digital Skills» annehmen und für die Ärzteschaft entsprechendes Informationsmaterial zur Verfügung stellen.

Fazit

Das Berufsbild der Ärztinnen und Ärzte wird sich durch KI verändern. In welcher Geschwindigkeit wird sich zeigen. Ärztinnen und Ärzte sollten diesen Prozess aktiv mitgestalten und sich diesem Wandel mit einer positiven Grundhaltung stellen. Dies nach dem Motto «KI for Future? – Ja, wir gestalten mit!»

Dank

Die FMH bedankt sich bei Prof. Dr. med. Michael Krauthammer, Prof. Dr. med. Nikola Biller-Andorno, Prof. Dr. iur. utr. Brigitte Tag für Ihre äusserst wertvollen Reviews und Beiträge. Prof. Jörg Goldhahn danken wir für die äusserst hilfreichen Informationen zum Wandel des ärztlichen Berufsbildes und Prof. Christian Johner für die äusserst hilfreiche Diskussion zu den Herausforderungen und Forderungen an KI in der Medizin. Prof. Roland Wiest danken wir für die wertvollen initialen Beiträge. Überaus grosser Dank geht an Herrn Walter Kern für die sprachliche Vereinfachung des Textes. Dem Atelier Richner danken wir herzlich für die hervorragende grafische Umsetzung. Den FMH Arbeitsgruppen danken wir für die wertvollen Hinweise und Diskussionen zu diesem Thema. Wir danken den der FMH angehörigen Ärzteorganisationen für ihre wertvollen Rückmeldungen zum Kapitel «Welche Forderungen stellt die FMH an die künstliche Intelligenz in der Medizin?».

Literatur

1. Densen, P. Challenges and opportunities facing medical education. *Trans. Am. Clin. Climatol. Assoc.* 122, (2011).
2. Liaw, W. & Kakadiaris, I. A. Artificial intelligence and family medicine: Better together. *Family Medicine* vol. 52 (2020).
3. Topol, E. J. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine* vol. 25 44–56 (2019).
4. NHS. The Topol Review. Preparing the healthcare workforce to deliver the digital future. An independent report on behalf of the Secretary of State for Health and Social Care. (2019).
5. FMH. Digital Trends Survey 2021. www.fmh.ch/digital-trends-survey-2021-de (2021).
6. OECD. Artificial Intelligence in Society. www.oecd.org/publications/artificial-intelligence-in-society-eedfee77-en.htm (2019).
7. Mccarthy, J. What is artificial intelligence? www-formal.stanford.edu/jmc (2007).
8. Winston, P. Artificial Intelligence. *Addison-Wesley, Reading, MA* (1992).
9. Society for Automotive Engineers. SAE J3016-Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. *SAE International* (2021).
10. ADAC. So funktioniert ein automatisiertes Auto. www.adac.de/rund-ums-fahrzeug/ausstattungs-technik-zubehoer/autonomes-fahren/grundlagen/auto-automatisiert-funktion (2020).
11. Russell, S. & Norvig, P. Artificial Intelligence A Modern Approach SECOND EDITION Prentice Hall Series in Artificial Intelligence (2003).
12. Domingos, P., Kok, S., Poon, H., Richardson, M. & Singla, P. Unifying Logical and Statistical AI. www.aaai.org/Papers/AAAI/2006/AAAI06-001.pdf
13. Molnar, C. Interpretable machine learning. A Guide for Making Black Box Models Explainable. christophm.github.io/interpretable-ml-book (2019).
14. Blarer, A., Buffat, M., Busch, C., Egloff, D., Fanzun, J., Haefliger, G., Langer, P., Loison, B., Luder, T., Malz, A., Scheidegger, E., Schneider, T., Schöll, M., Schwaar, P., Stämpfli, M., Weber, V. Herausforderungen der künstlichen Intelligenz. Bericht der interdepartementalen Arbeitsgruppe «Künstliche Intelligenz» an den Bundesrat (2019).
15. Taddy, M. & Booth, C. The Technological Elements of Artificial Intelligence. www.nber.org/papers/w24301 (2018).
16. Krzyk, K. Towards Data Science. Coding Deep Learning For Beginners. towardsdatascience.com/coding-deep-learning-for-beginners-types-of-machine-learning-b9e651e1ed9d (2018).
17. IBM Cloud Education. Supervised vs. Unsupervised Learning: What's the Difference? www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning (2015).
18. IBM Cloud Education. Machine Learning. <https://www.ibm.com/cloud/learn/machine-learning> (2020).
19. Noorbakhsh-Sabet, N., Zand, R., Zhang, Y. & Abedi, V. Artificial Intelligence Transforms the Future of Health Care. *American Journal of Medicine* vol. 132 (2019).
20. Johner, C. Künstliche Intelligenz in der Medizin. www.johner-institut.de/blog/regulatory-affairs/kuenstliche-intelligenz-in-der-medizin/#section_scroll2 (2020).
21. Dalton, L., Ballarin, V. & Brun, M. Clustering Algorithms: On Learning, Validation, Performance, and Applications to Genomics. *Curr. Genomics* 10, (2009).
22. Komorowski, M., Celi, L. A., Badawi, O., Gordon, A. C. & Faisal, A. A. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nat. Med.* 24 (2018).
23. Merriam-Webster. Algorithm. *Merriam-Webster.com dictionary*. www.merriam-webster.com/dictionary/algorithm.
24. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks (2017). <https://dl.acm.org/doi/10.1145/3065386>.
25. Rosenblatt, F. The Perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review* vol. 65 (1958).
26. Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. *Neural Comput.* 2006 Jul;18(7):1527-54.
27. Merriam-Webster. Neural Network. <https://www.merriam-webster.com/dictionary/neural-network>.
28. IBM Cloud Education. Neural Networks. www.ibm.com/cloud/learn/neural-networks#toc-neural-net-u3voPJVU (2020).
29. IBM Cloud Education. Deep Learning. <https://www.ibm.com/cloud/learn/deep-learning> (2020).
30. Pereira T, Tran N, Gadhoumi K, Pelter MM, Do DH, Lee RJ, Colorado R, Meisel K, Hu X. Photoplethysmography based atrial fibrillation detection: a review. *NPJ Digit Med.* 2020 Jan 10;3:3. doi: 10.1038/s41746-019-0207-9.
31. Abràmoff, M. D., Lavin, P. T., Birch, M., Shah, N. & Folk, J. C. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *Digit. Med.* 1, 39 (2018).
32. Yu Kun-Hsing, Beam Andrew L & Kohane Isaac S. Artificial intelligence in healthcare. *Nat. Biomed. Eng.* 2, 719–731 (2018).
33. Keane, P. A. & Topol, E. J. With an eye to AI and autonomous diagnosis. *npj Digit. Med.* 1, (2018).
34. Rohani, N. & Eslahchi, C. Drug-Drug Interaction Predicting by Neural Network Using Integrated Similarity. *Sci. Rep.* 9, (2019).
35. Raja, K., Patrick, M., Elder, J. T. & Tsoi, L. C. Machine learning workflow to enhance predictions of Adverse Drug Reactions (ADRs) through drug-gene interactions: Application to drugs for cutaneous diseases. *Sci. Rep.* 7, (2017).

36. Hyland SL, Faltys M, Hüser M, Lyu X, Gumbsch T, Esteban C, Bock C, Horn M, Moor M, Rieck B, Zimmermann M, Bodenham D, Borgwardt K, Rättsch G, Merz TM. Early prediction of circulatory failure in the intensive care unit using machine learning. *Nat. Med.* 2020 Mar;26(3):364-373. doi: 10.1038/s41591-020-0789-4.
37. Stehlik J, Schmalfluss C, Bozkurt B, Nativi-Nicolau J, Wohlfahrt P, Wegerich S, Rose K, Ray R, Schofield R, Deswal A, Sekaric J, Anand S, Richards D, Hanson H, Pipke M, Pham M. Continuous Wearable Monitoring Analytics Predict Heart Failure Hospitalization: The LINK-HF Multicenter Study. *Circ Heart Fail.* 2020 Mar;13(3):e006513. doi: 10.1161/CIRCHEARTFAILURE.119.006513.
38. Trezzini, B., Meyer, B., Ivankovic, M., Jans, C. & Golder, L. Der administrative Aufwand der Ärzteschaft nimmt weiter zu. *Schweizerische Ärztezeitung* (2020). doi: 10.4414/saez.2020.18482.
39. Rajkomar, A., Dean, J. & Kohane, I. Machine Learning in Medicine. *N. Engl. J. Med.* 380, (2019).
40. Coiera, E., Kocaballi, B., Halamaka, J. & Laranjo, L. The digital scribe. *npj Digit. Med.* 1, (2018).
41. Chiu, C.-C., Tripathi, A., Chou, K., Co, C., Jaitly, N., Jaunzeikare, D., Kannan, A., Nguyen, P., Sak, H., Sankar, A., Tansuwan, J., Wan, N., Wu, Y., Zhang, X. (2018) Speech Recognition for Medical Conversations. *Proc. Interspeech 2018*, 2972-2976, doi: 10.21437/Interspeech.2018-40.
42. Jing, B., Xie, P. & Xing, E. P. On the automatic generation of medical imaging reports. in *ACL 2018 - 56th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers) vol. 1* (2018).
43. Gehlbach, P. L. Robotic surgery for the eye. *Nature Biomedical Engineering* vol. 2 (2018).
44. EU. Bericht über die Auswirkungen künstlicher Intelligenz, des Internets der Dinge und der Robotik in Hinblick auf Sicherheit und Haftung. (2020) doi: 10.1787/ab757416-en.
45. Unabhängige hochrangige Expertengruppe für Künstliche Intelligenz eingesetzt von der Europäischen Kommission. Ethik-Leitlinien für eine vertrauenswürdige KI. digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (2018).
46. Fraunhofer IAIS. Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz – KI Prüfkatalog (2021). www.iais.fraunhofer.de/ki-pruefkatalog.
47. Fraunhofer IAIS. Vertrauenswürdiger Einsatz von künstlicher Intelligenz. (2019).
48. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential Biases in Machine Learning Algorithms using Electronic Health Record Data. *JAMA Intern Med.* 2018 Nov 1;178(11):1544-1547. doi: 10.1001/jamainternmed.2018.3763.
49. Rajkomar, A., Hardt, M., Howell, M.D., Chin, M.H. & Author, C. Ensuring Fairness in Machine Learning to Advance Health Equity HHS Public Access. *Ann Intern Med* 169, 866–872 (2018).
50. Lyell, D. & Coiera, E. Automation bias and verification complexity: a systematic review (2016). doi: 10.1093/jamia/ocw105.
51. Tramèr F., Atlidakis V., Geambasu R., Hsu D., Hubaux J.P., Humbert M., Juels A., Lin H. FairTest: Discovering Unwarranted Associations in Data-Driven Applications. *Proceedings - 2nd IEEE European Symposium on Security and Privacy, EuroS and P* (2017). doi: 10.1109/EuroSP.2017.29.
52. FairTest. <https://github.com/columbia/fairtest>.
53. Nothwang, W.D., McCourt, M.J., Robinson, R.M., Burden, S. A. & Curtis, J. W. The human should be part of the control loop? in *Proceedings - 2016 Resilience Week, RWS* (2016). doi: 10.1109/RWEEK.2016.7573336.
54. Campanella G, Hanna MG, Geneslaw L, Mirafior A, Werneck Krauss Silva V, Busam KJ, Brogi E, Reuter VE, Klimstra DS, Fuchs TJ. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med.* 2019 Aug;25(8):1301-1309. doi: 10.1038/s41591-019-0508-1.
55. Benjamins, S., Dhunoo, P. & Meskó, B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *npj Digit. Med.* 3, (2020).
56. FDA. Evaluation of automatic class III designation for Osteodetect. www.fda.gov/about-fda/cdrh-transparency/evaluation-automatic-class-iii-designation-de-novo-summaries.
57. FDA. Critical Care Suite: Radiological Computer Assisted Triage and Notification Software. www.accessdata.fda.gov/cdrh_docs/pdf18/K183182.pdf
58. Narang A, Bae R, Hong H, Thomas Y, Surette S, Cadieu C, Chaudhry A, Martin RP, McCarthy PM, Rubenson DS, Goldstein S, Little SH, Lang RM, Weissman NJ, Thomas JD. Utility of a Deep-Learning Algorithm to Guide Novices to Acquire Echocardiograms for Limited Diagnostic Use. *JAMA Cardiol.* 2021 Jun 1;6(6):624-632. doi: 10.1001/jamacardio.2021.0185.
59. Yang, A., Wang, F. & Sheu, Y.-H. Illuminating the Black Box: Interpreting Deep Neural Network Models for Psychiatric Research. *Front. Psychiatry* www.frontiersin.org 11, 551299 (2020).
60. Collins, G.S., Reitsma, J.B., Altman, D.G. & Moons, K.G.M. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. *Eur. Urol.* 67, 1142–1151 (2015).
61. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, Arnaout R, Kohane IS, Saria S, Topol E, Obermeyer Z, Yu B, Butte AJ. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. *Nat Med.* 2020 Sep;26(9):1320-1324. doi: 10.1038/s41591-020-1041-y.
62. Mongan, J., Moy, L. & Kahn, C. E. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. *Radiol. Artif. Intell.* 2, e200029 (2020).
63. Liu, X., Cruz Rivera, S., Moher, D., Calvert, M. J. & Denniston, A. K. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nature medicine* vol. 26 (2020).

64. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group; SPIRIT-AI and CONSORT-AI Steering Group; SPIRIT-AI and CONSORT-AI Consensus Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat Med*. 2020 Sep;26(9):1351-1363. doi: 10.1038/s41591-020-1037-7.
65. Fedlex. Bundesgesetz über die Produktheftpflicht, SR 221.112.944. (2010).
66. Fedlex. Bundesgesetz über die Produktesicherheit, SR 930.11. (2010).
67. Landolt, H. & Herzog-Zwitter, I. *Arzt Haftungsrecht*. in 359ff (Dike Verlag, 2014).
68. Regina E. Aebi-Müller, Walter Fellmann, Thomas Gächter, Bernhard Rüttsche & Brigitte Tag. *Arztrecht*. in 267–328 (Stämpfli Bern, 2016).
69. Song X, Yu ASL, Kellum JA, Waitman LR, Matheny ME, Simpson SQ, Hu Y, Liu M. Cross-site transportability of an explainable artificial intelligence model for acute kidney injury prediction. *Nat Commun*. 2020 Nov 9;11(1):5668. doi: 10.1038/s41467-020-19551-w.
70. Lacaze P, Pinese M, Kaplan W, Stone A, Brion MJ, Woods RL, McNamara M, McNeil JJ, Dinger ME, Thomas DM. The Medical Genome Reference Bank: a whole-genome data resource of 4000 healthy elderly individuals. Rationale and cohort design. *Eur J Hum Genet*. 2019 Feb;27(2):308-316. doi: 10.1038/s41431-018-0279-z.
71. Rose, S. Machine Learning for Prediction in Electronic Health Data. *JAMA Network Open* vol. 1 (2018).
72. Nevin, L. Advancing the beneficial use of machine learning in health care and medicine: Toward a community understanding. *PLoS Medicine* vol. 15 (2018).
73. Fryback, D. G. & Thornbury, J. R. The efficacy of diagnostic imaging. *Med. Decis. Mak.* 11, (1991).
74. INFANT Collaborative Group. Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *Lancet*. 2017 Apr 29;389(10080):1719-1729. doi: 10.1016/S0140-6736(17)30568-8.
75. Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., et al. The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. (2018). <https://doi.org/10.17863/CAM.22520>.
76. Rieke N, Hancox J, Li W, Milletari F, Roth HR, Albarqouni S, Bakas S, Galtier MN, Landman BA, Maier-Hein K, Ourselin S, Sheller M, Summers RM, Trask A, Xu D, Baust M, Cardoso MJ. The future of digital health with federated learning. *NPJ Digit Med*. 2020 Sep 14;3:119. doi: 10.1038/s41746-020-00323-1.
77. Warnat-Herresthal S, Schultze H, Shastry KL, Manmohan S, Mukherjee S, Garg V, Sarveswara R, Händler K, Pickkers P, Aziz NA, Ktena S, Tran F, Bitzer M, Ossowski S, Casadei N, Herr C, Petersheim D, Behrends U, Kern F, Fehlmann T, Schommers P, Lehmann C, Augustin M, Rybniker J, Altmüller J, Mishra N, Bernardes JP, Krämer B, Bonaguro L, Schulte-Schrepping J, De Domenico E, Siever C, Kraut M, Desai M, Monnet B, Saridaki M, Siegel CM, Drews A, Nuesch-Germano M, Theis H, Heyckendorf J, Schreiber S, Kim-Hellmuth S; COVID-19 Aachen Study (COVAS), Nattermann J, Skowasch D, Kurth I, Keller A, Bals R, Nürnberg P, Riess O, Rosenstiel P, Netea MG, Theis F, Mukherjee S, Backes M, Aschenbrenner AC, Ulas T; Deutsche COVID-19 Omics Initiative (DeCOI), Breteler MMB, Giamarellos-Bourboulis EJ, Kox M, Becker M, Cheran S, Woodacre MS, Goh EL, Schultze JL. Swarm Learning for decentralized and confidential clinical machine learning. *Nature*. 2021 Jun;594(7862):265-270. doi: 10.1038/s41586-021-03583-3.
78. WIPO. Artificial Intelligence and Intellectual Property. *World Intellectual Property Organization*. www.wipo.int/about-ip/en/artificial_intelligence/policy.html
79. Hilty, R. M. *Urheberrecht*. (Stämpfli juristische Lehrbücher, 2020).
80. Vieth, A. Tom L. Beauchamp, James F. Childress (2009) *Principles of Biomedical Ethics*. 6. Auflage. *Ethik der Medizin* 22, (2010).
81. Europäisches Parlament. Regelung der zivilrechtlichen Haftung beim Einsatz künstlicher Intelligenz. www.europarl.europa.eu/doceo/document/TA-9-2020-0276_DE.html.
82. Gless, S. Von der Verantwortung einer E-Person. *Goltdammer's Arch. für Strafr.* 164, 324–329 (2017).
83. FMH. Standesordnung der FMH. www.fmh.ch/ueber-die-fmh/statuten-reglemente.cfm#112408 (2020).
84. FDA. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)-Discussion Paper and Request for Feedback. www.fda.gov/media/122535/download (2019).
85. FDA. Software as a Medical Device (SaMD) Action Plan. www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device (2021).
86. FDA. Good Machine Learning Practice for Medical Device Development: Guiding Principles. www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles (2021).
87. EU. Verordnung des Europäischen Parlaments und Rates. Zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz (Gesetz über Künstliche Intelligenz) und Änderung bestimmter Rechtsakte der Union. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206> (2021).

88. Council of the European Union. Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- Presidency compromise text, 14278-21. (2021).
89. EU. Verordnung (EU) 2017/745 des Europäischen Parlaments und Rates über Medizinprodukte, zur Änderung der Richtlinie 2001/83/EG, der Verordnung (EG) Nr. 178/2002 und der Verordnung (EG) Nr. 1223/2009 und zur Aufhebung der Richtlinien 90/385/EWG und 93/42/EWG des Rates. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A32017R0745> (2017).
90. EU. Neue Vorschriften für künstliche Intelligenz – Fragen und Antworten (2021). https://ec.europa.eu/commission/presscorner/detail/de/QANDA_21_1683
91. Karner, E., Koch, B. A. & Geistfeld, M. A. Comparative Law Study on Civil Liability for Artificial Intelligence Justice and Consumers. (2020).
92. Johner, C. Regulatorische Anforderungen an Medizinprodukte mit Machine Learning. www.johner-institut.de/blog/iec-62304-medizinische-software/regulatorische-anforderungen-an-medizinprodukte-mit-machine-learning (2021).
93. Johner, C., Molnar, C., Purde, A., Rad, A., Dierks, C., Bunk, S., Piechottka, S. Leitfaden zur KI bei Medizinprodukten. https://github.com/johner-institut/ai-guideline/blob/master/Guideline-AI-Medical-Devices_DE.md#leitfaden-zur-ki-bei-medizinprodukten.
94. ICMRA. Informal Innovation Network Horizon Scanning Assessment Report – Artificial Intelligence. www.icmra.info/drupal/en/strategicinitiatives/innovation (2021).
95. Digital Society Initiative. Ein Rechtsrahmen für Künstliche Intelligenz. www.dsi.uzh.ch/de/research/projects/strategy-lab/strategy-lab-21.html (2021).
96. Graf, F. & Obrecht, L. Rechtliche Rahmenbedingungen für Künstliche Intelligenz in der Schweiz. *Jusletter* (2021) doi: 10.38023/431d79f8-7b98-4828-97be-b346864f371d.
97. Fischer, J. & Fuchs, M. Brauchen wir eine Legaldefinition für künstliche Intelligenz? *Jusletter* (2021) doi: 10.38023/681145f8-fb8b-4354-90d7-79a7b5112fdc.
98. Fedlex. Medizinprodukteverordnung. SR 812.213. www.fedlex.admin.ch/eli/cc/2020/552/de (2020).
99. EU. Verordnung 2017/746 des Europäischen Parlaments und des Rates über In-vitro-Diagnostika und zur Aufhebung der Richtlinie 98/79/EG und des Beschlusses 2010/227/EU der Kommission. <https://eur-lex.europa.eu/legal-content/de/ALL/?uri=CELEX%3A32017R0746> (2017).
100. American Medical Association. Augmented Intelligence in Health Care policy report. (2019).
101. Crochane Deutschland. Evidenzbasierte Medizin. www.cochrane.de/ueber-uns/evidenzbasierte-medizin.
102. WHO. Ethics and governance of artificial intelligence for health. <https://www.who.int/publications/i/item/9789240029200>. (2021).
103. Bonvin R., Buhmann J., Cotrini Jimenez C., Egger M., Geissler A., Krauthammer M., Schirlo C., Spiess C., Steurer J., Vokinger K. N., Vogt J. Neue Lernziele für das Medizinstudium erarbeitet: Studierende auf den Einsatz von maschinellem Lernen vorbereiten. In: *Schweizerische Ärztezeitung* (2022), 4, S. 98-101.
104. Buhmann J., Felix J., Gächter T., Kowatsch T., Lehmann R., von Lutterotti N., Schedler K., Steurer J., Wolfrum C. (2018) Digitalisierung der Medizin: Konsequenzen für die Ausbildung. *Schweizerische Ärztezeitung*, 99 (42). 1441-1444.
105. Schüller, K., Busch, P. & Hindinger, C. Future Skills: Ein Framework für Data Literacy- Kompetenzrahmen und Forschungsbericht. (2019).
106. Misra, A., Jütting, J. & Kuonen, D. Advancing data literacy in the post-pandemic world. (2021).
107. CPRD. Clinical Practice Research Datalink. www.cprd.com.
108. CPME. CPME Policy on AI in Health Care. www.cpme.eu (2019).
109. American Medical Association. Augmented intelligence in health care: Policy. www.ama-assn.org/system/files/2019-08/ai-2018-board-policy-summary.pdf (2018).
110. IRGC. The Governance of Decision-Making Algorithms. www.epfl.ch/research/domains/irgc/specific-risk-domains/projects-cybersecurity/the-governance-of-decision-making-algorithms (2018).
111. SAMW. Expert workshop report organised by SATW and SAMW on 'Artificial Intelligence in medicine - can it be trusted?' www.assm.ch/de/Projekte/Uebersicht-der-Projekte/Kuenstliche-Intelligenz.html (2018).
112. Frank, J.R., Snell, L., Sherbino, J. & Royal College of Physicians and Surgeons of Canada. CanMEDS 2015 physician competency framework (2015).
113. Rampton, V., Mittelman, M. & Goldhahn, J. Implications of artificial intelligence for medical education. *The Lancet Digital Health* vol. 2 (2020).
114. Wartman, S.A. & Donald Combs, C. Medical education must move from the information age to the age of artificial intelligence. *Acad. Med.* 93, (2018).
115. Weissmann, Y., Wolfrum, C. & Goldhahn, J. Bachelor Humanmedizin an der ETH Zürich. *Praxis (Bern. 1994)*. 109, (2020).
116. CAIM. Master of Science Artificial Intelligence in Medicine. www.caim.unibe.ch/education/msc_aim/index_eng.html (2021).

FMH – Verbindung der Schweizer
Ärztinnen und Ärzte
Postfach
3000 Bern 16

Tel. 031 359 11 11
info@fmh.ch

