Forderungen an die Künstliche Intelligenz (KI) aus Sicht der FMH

Version 1.0 vom 21. Februar 2022

Mit Hilfe der KI sollen alle Prozesse entlang des Patientenbehandlungsweges sinnvoll und zweckmässig weiterentwickelt und verbessert werden. Dabei gibt es einige Herausforderungen, die der Einsatz von KI in der Medizin mit sich bringt, zu beachten.

Die KI-Systeme müssen einen medizinischen Mehrwert schaffen. Sie sollen anhand klinischer Settings validiert sein und auch nach der Inbetriebnahme regelmässig überprüft werden. Dafür sind grosse annotierte Datensätze, welche die jeweilige Patientenpopulation sowie das Krankheitsbild korrekt abbilden, unabdingbar. Zudem braucht es Referenzdatensätze von Gesunden, eine robuste Validierung und die Integration in den klinischen Alltag. Auch die Transparenz der verwendeten Methodik und der verwendeten Datensätze zu Trainings-, Test-, und Validierungszwecken muss gegeben und möglichst international standardisiert sein.

Zudem ist auf mögliche Verzerrungen (Bias) in den Daten, dem Design, den Interaktionen zwischen Mensch und Maschine zu testen und diese sind auszuweisen. Von Bedeutung ist ein geeignetes Qualitätsmanagement ebenso wie Qualitätsanforderungen an die Datensammlung. Ärztinnen und Ärzte müssen zwingend Entscheidungsträger bleiben und die Möglichkeit haben, die Entscheidungen des KI-Systems zu übersteuern. Essenziell für den Erfolg und die Akzeptanz des Einsatzes von KI-Systemen in der Medizin ist, dass in deren Entwicklung nachweislich Ärztinnen und Ärzte sowie Patientinnen und Patienten einbezogen werden.

Die Forderungen der FMH an KI-Systeme beziehen sich auf diejenigen Systeme, die zu diagnostischen und/oder therapeutischen Zwecken herangezogen werden und somit das Potential haben, ärztliche Kerndienstleistungen zu ersetzen. Im Folgenden werden KI-Systeme betrachtet, die nach Klasse IIA und höher gemäss Medizinprodukteverordnung (Medical Device Regulation, MDR) eingestuft werden. Bei der Erarbeitung der Forderungen wurden einzelne Punkte bestehender oder vorgeschlagener Gesetze, Regularien und Leitfäden (MDR [1], MepV [2], IVDR [3], Gesetz über KI [4], Leitfaden zur KI bei Medizinprodukten [5], Guiding Principles for Good Machine Learning Practice for Medical Device Development [6]) berücksichtigt, die aus Sicht der Ärzteschaft wichtig sind.

Die Forderungen an die KI-Systeme unter Berücksichtigung der aktuellen rechtlichen und technischen Situation, lauten im Einzelnen:

1. Das KI-System soll die menschliche Intelligenz und die Beziehung zwischen Arzt und Patient stärken und nicht ersetzen[7].
2. Das KI-System muss sich an den Prinzipien der evidenzbasierten Medizin (EbM) orientieren. EbM ist der gewissenhafte, ausdrückliche und vernünftige Gebrauch der gegenwärtig besten externen, wissenschaftlichen Evidenz für Entscheidungen in der medizinischen Versorgung individueller Patienten. Die Praxis der EbM bedeutet die Integration individueller klinischer Expertise mit der bestverfügbaren externen Evidenz aus systematischer Forschung [8].
3. Die Leistungen eines KI-Systems müssen wirkungsvoll und nutzbringend sein. Die Leistungen des KI-Systems müssen daraufhin periodisch überprüft werden.
4. Das KI-System muss nach der Inbetriebnahme im medizinischen Setting regelmässig für die Patientenpopulation, auf die es zum erklärten Zweck angewendet wird, überprüft werden (Post-Market-Surveillance). Es sind unverzüglich Korrekturen vorzunehmen, wenn sie sich aufgrund der Überwachung als geboten erweisen.

5. Nach dem gegenwärtigen Stand der rechtlichen Situation bleiben Ärztinnen und Ärzte zusammen mit ihren Patientinnen und Patienten Entscheidungsträger in Bezug auf den Einsatz und den Umgang mit den Empfehlungen des KI-Systems.
6. Neben einer angemessenen Schulung der Ärztinnen und Ärzte sowie weiteren involvierten Gesundheitsfachpersonen im Umgang mit dem KI-System muss ihnen eine «Gebrauchsanweisung» beiliegen, die klar aufzeigt:

    a. Welche Rolle der Ärztin oder dem Arzt beim Betrieb des Systems zugeschrieben wird. Als Beispiel sei hier die Entscheidungshoheit bei «in the loop»- oder «on the loop»-Systemen genannt.
    b. Für welchen Zweck bzw. für welches Krankheitsbild das KI-System eingesetzt werden darf und für welchen Anwendungsfall nicht.
    c. Wie mit dem KI-System umzugehen ist bzw. wie es zu verwenden ist. Weiterhin muss aufgezeigt sein, welche Daten wie eingegeben werden müssen, wie Fehlfunktionen und besondere Vorfälle erkannt werden können, woran erkannt werden kann, dass das System einwandfrei funktioniert, sowie das System, falls sinnvoll, übersteuert werden kann.
    d. Auf welcher Methodik das KI-System basiert.
    e. Wie das KI-System getestet, trainiert und validiert wurde. Die jeweils verwendeten unabhängigen Datensätze und deren Charakteristika müssen angegeben sein (Grösse, Verteilung Frau, Mann, Alter etc.). Deren Eignung (aus statistischer und medizinischer/klinischer Perspektive) inklusive der Beschreibung der Tests auf mögliche Verzerrungen (Bias) und Fehler, sowie die Kennzahlen zur Genauigkeit der Leistung und der klinischen Wirksamkeit des KI-Systems, auch im Vergleich zu bestmöglichen alternativen Methoden und KI-Modellen, müssen beschrieben sein.
    f. Wie das KI-System den Schutz und die Sicherheit der Daten und die Informationssicherheit garantiert und insbesondere wo die Daten (Personendaten und Metadaten) und in welcher Form diese gespeichert werden.
    g. An wen sich der Nutzer im Falle von Fragen und Problemen, welche die Funktion und die Sicherheit (z. B. Datensicherheit) betreffen, wenden kann.

    Die «Gebrauchsanweisung» für die Nutzer muss alle erforderlichen Informationen zur einwandfreien Verwendung des KI-Systems beinhalten. Es muss sichergestellt werden, dass dadurch Ärztinnen und Ärzten ihren Informations- und Sorgfaltspflichten nachkommen können. Die Anwender der KI müssen durch die «Gebrauchsanweisung» in die Lage versetzt werden können, den Betrieb des KI-Systems zu verstehen, zu überwachen und bei Bedarf auch zu übersteuern.

7. Bei der Entwicklung von KI-Systemen müssen Ärztinnen und Ärzte sowie Patientinnen und Patienten für die Erhebung der Nutzungsanforderungen einbezogen werden. Damit ihre Nutzungsziele optimal und effizient erreicht werden können, muss insbesondere spezifiziert werden, 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.

8. Ärztinnen und Ärzte müssen über den Gebrauch von KI-Systemen, die ihre Arbeit indirekt beeinflussen, informiert werden. Denkbar sind Systeme, die im Hintergrund die Qualität der ärztlichen Arbeit überwachen, oder Systeme, die Kostengutsprachen für Medikamente übernehmen.

KI-Systeme müssen auf die lokale Datensituation angepasst werden. Dazu braucht es eine nationale Dateninfrastruktur, die es ermöglicht die KI-Systeme mit grossen qualitativ hochwertigen annotierten medizinischen/klinischen Datensätzen zu testen, zu trainieren und die Reproduzierbarkeit der Ergebnisse durch eine robuste Validierung sicherzustellen. Die Datensätze müssen die jeweilige zu testende lokale Patientenpopulation repräsentieren und unter Beachtung des Datenschutzes und der Datensicherheit erhoben, verarbeitet und gespeichert werden. Wichtig ist hier insbesondere ein zertifiziertes Qualitätsmanagement und ein zertifiziertes Labeling (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, welche ein dezentrales Lernen von Daten ermöglicht.

Die europäische Ärztegesellschaft (CPME, Standing Committee of European Doctors) und die amerikanische Ärztegesellschaft (AMA, American Medical Association) haben ebenfalls minimale Forderungen an KI-Systeme definiert [9, 10]. Workshops zu den Themen «Regulierung von KI getriebenen Algorithmen zur Entscheidungsfindung» (IRGC, International Risk Governance Center) und «KI in der Medizin – kann man ihr vertrauen» (SAMW, Schweizerische Akademie der medizinischen Wissenschaften; SATW, Schweizerische Akademie der technischen Wissenschaften) fanden auch in der Schweiz statt [11, 12].

Glossar

Begriff	Erklärung
«in the loop»	Eines von möglichen Szenarien, welches für die Zusammenarbeit zwischen Mensch und Maschine denkbar ist. Beim «in the loop» Szenario ist der Mensch für den Arbeitsprozess unabdingbar. Das bedeutet die Maschine kann keine Aufgabe ohne die Bestätigung durch den Menschen abschliessen. Zudem kann der Mensch Entscheidungen, die vom KI-System selbständig getroffen werden korrigieren, überschreiben und kompensieren [13].
«on the loop»	Eines von möglichen Szenarien, welches für die Zusammenarbeit zwischen Mensch und Maschine denkbar ist. Beim «on the loop» Szenario erledigen Maschinen 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 Entscheidungen, die vom KI-System selbständig getroffen werden nachträglich korrigieren, überschreiben und kompensieren [13].
föderiertes Lernen	Das Zusammenführen von Personendaten zu grossen Datensätzen, sogenannte «Big Data» Datensätze stellt aus Sicht des Datenschutzes eine Herausforderung dar, die das Teilen der Daten begrenzt. KI-Systeme benötigen allerdings für das Trainieren ihrer Modelle grosse für die Patientenpopulation relevante Datensätze. Neue technische Ansätze, die es den KI-Systemen erlauben, auf verteilte Datensätze zuzugreifen, ohne dass diese physisch zusammengelegt werden müssen, bieten hierfür eine Lösung. Ein Beispiel einer solchen konkreten Technologie ist das föderierte Lernen [14]. Beim föderierten Lernen werden die KI-Modelle in einem ersten Schritt auf verschiedenen lokalen Spitaldaten trainiert. In einem zweiten Schritt werden die lokalen Modellparameter in einem globalen Modell zusammengeführt.

[h] Als Beispiel kann die UK genannt werden, welche im CPRD Netzwerk Daten von Hausärzten sammelt [15].

Literatur

1. 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).
2. Fedlex. *Medizinprodukteverordnung. SR 812.213.* https://www.fedlex.admin.ch/eli/cc/2020/552/de (2020).
3. 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).
4. 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).
5. Johner, C. . *et al.* 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*.
6. FDA. *FDA Issues Guiding Principles for Good Machine Learning Practice for Medical Device Development*. https://www.lifesciencesperspectives.com/2021/11/04/fda-issues-guiding-principles-for-good-machine-learning-practice-for-medical-device-development/ (2021).
7. American Medical Association. *Augmented Intelligence in Health Care policy report*. (2019).
8. Crochane Deutschland. Evidenzbasierte Medizin. https://www.cochrane.de/ebm.
9. CPME. *CPME Policy on AI in Health Care*. www.cpme.eu (2019).
10. American Medical Association. *Augmented intelligence in health care: Policy*. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwim6uLc1tv1AhVdg_0HHZ0LC9UQFnoECAkQAQ&url=https%3A%2F%2Fwww.ama-assn.org%2Fsystem%2Ffiles%2F2019-08%2Fai-2018-board-policy-summary.pdf&usg=AOvVaw2lCcUxas5_VzBd2PXyDPjI (2018).
11. IRGC. The Governance of Decision-Making Algorithms. *https://www.epfl.ch/research/domains/irgc/specific-risk-domains/projects-cybersecurity/the-governance-of-decision-making-algorithms/* (2018).
12. SAMW. *Expert workshop report organised by SATW and SAMW on 'Artificial Intelligence in medicine - can it be trusted?'* https://www.assm.ch/de/Projekte/Uebersicht-der-Projekte/Kuenstliche-Intelligenz.html (2018).
13. Fraunhofer IAIS. *Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz - KI Prüfkatalog*. www.iais.fraunhofer.de/ki-pruefkatalog.
14. Rieke, N. *et al.* PERSPECTIVE The future of digital health with federated learning. doi:10.1038/s41746-020-00323-1.
15. CPRD. *Clinical Practice Research Datalink*. https://www.cprd.com/.