

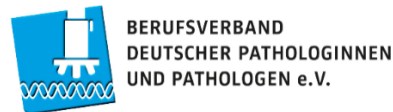
Revision von MDR und IVDR

Kommentierung und Vorschläge
aus der Sicht der akademischen Forschung
und akademischen Verbundforschung

gemeinsam erstellt und unterstützt von:



fit4translation





1 Präambel

Die Verordnung (EU) 2017/745 über Medizinprodukte (MDR) sowie die Verordnung (EU) 2017/746 über In-vitro-Diagnostika (IVDR) wurden eingeführt, um Patientensicherheit zu erhöhen, Innovation zu fördern und einheitliche Standards im europäischen Binnenmarkt zu schaffen. In der Praxis zeigt sich jedoch, dass die aktuellen Regelungen in wesentlichen Punkten ihr Ziel verfehlen und stattdessen neue Hürden errichten – insbesondere für die translationale Forschung in der Universitätsmedizin und für die akademische Verbundforschung und für die Anwendung in der medizinischen Versorgung.

MDR und IVDR stellen hohe Anforderungen an die Sicherheit und Leistung von Medizinprodukten und In-vitro Diagnostika, die in die klinische Forschung oder die klinische Anwendung gebracht werden sollen. Diese sind jedoch häufig mit einem hohen bürokratischen Aufwand verbunden, der in keinem angemessenen Verhältnis zu den tatsächlichen Risiken für Patient*innen, Anwender*innen, Betreibern und Datengebern der Produkte steht. Insbesondere sind viele Regelungen nicht für digitale Medizinprodukte wie Software und KI-Systeme anwendbar, bzw. stellen umgekehrt Software und KI-Systeme das regulatorische System vor neue Herausforderungen. An Universitätskliniken, in Forschungsverbünden und in öffentlich geförderten Projekten führt diese Überlastung dazu, dass klinisch notwendige Innovationen verzögert oder verhindert werden. Aber auch die Einführung in die Routineversorgung kann durch die Anforderungen unangemessen erschwert werden. Gleichzeitig entsteht ein internationaler Wettbewerbsnachteil, da in anderen Regionen – etwa den USA – flexiblere, risikoadäquatere Regulierungsmodelle etabliert sind.

In dem vom deutschen Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) geförderten Begleitforschungsvorhaben zur Medizininformatik-Initiative wurden in Zusammenarbeit mit der TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e. V., unter Einbeziehung des Medizinischen Fakultätentags und des Verbands der Universitätsklinika Deutschlands e. V. sowie weiterer Fachgesellschaften und Verbände (aufgeführt unter Abschnitt 15) konkrete Kommentierungen und Vorschläge erarbeitet, die darauf abzielen, die MDR/IVDR praxisgerechter auszugestalten. Dabei stehen drei Anliegen im Vordergrund:

- **Patientensicherheit** gewährleisten durch eine risikoadäquate Regulierung, die die besonderen Eigenschaften von Anwendungsfall, Software, KI (künstlicher Intelligenz) und innovativen Versorgungsmodellen berücksichtigt.
- **Translationale Forschung** ermöglichen durch Entlastung von übermäßiger Bürokratie, damit neue diagnostische und therapeutische Ansätze schneller aus der Forschung in die klinische Anwendung gelangen.
- **Regulatorische Kohärenz** und damit herstellen durch eine Harmonisierung mit angrenzenden Rechtsakten (insbesondere Artificial Intelligence - AI-Act, NIS2-Richtlinie, DSGVO, EHDS-Verordnung) und die Vermeidung von Doppelregulierungen.

Die nachfolgenden Kommentierungen sollen eine Grundlage bieten, um die regulatorischen Rahmenbedingungen weiterzuentwickeln – hin zu einem System, das Patientensicherheit mit Innovationsfähigkeit und Forschungsfreiheit in Einklang bringt. Zur Stärkung von Translation und Transfer.



2 Inhalt

1	Präambel	2
2	Inhalt	3
3	Nicht risikoangepasste Klassifizierung von Software (Anhang VIII Regel 11)	5
3.1	Ist (aktuelle Regelung):	5
3.2	Problem	5
3.3	Vorschlag für eine Regelung:	5
4	Klinische Bewertung	7
4.1	Antizipierende klinische Bewertung und Stärkung der Überwachung im Feld	7
4.2	Anpassung von Studiendesign Voraussetzungen an Risiko-Klassifizierung bei klinischen Studien und Leistungsprüfungen von Software, digitalen Biomarkern und KI-Systemen	9
5	Klinische Prüfungen – Einwilligung bei Cluster Randomisierten Studien	10
5.1	Ist (aktuelle Regelung)	10
5.2	Problem	10
5.3	Vorgeschlagene Regelung	11
6	Eigenherstellung nach Art. 5(5) MDR	11
6.1	Gewinnung von „Erfahrungen“ zum Einsatz von Eigenherstellungen	11
6.2	Einheitliche Regelung für Hochschulmedizin im Kooperationsmodell / Klinikverbund, Praxisnetzwerke und öffentlich geförderte Verbundforschungsprojekte	13
6.3	Abgabe als Open Source Software	14
7	Eigenherstellungen nach IVDR	14
8	Öffentlich zugängliche Informationen über medizinische KI-Systeme	14
8.1	Ist (aktuelle Regelung)	14
8.2	Problem:	15
8.3	Vorschlag für eine Regelung	15
9	Open Source Software und SOUP	16
9.1	Ist (Aktuelle Regelung)	16
9.2	Problem	16
9.3	Vorgeschlagene Regelung	16
10	Harmonisierung der MDR / IVDR mit anderen Richtlinien und Verordnungen	17
10.1	AI-Act	17
10.2	Cybersecurity (NIS2, KRITIS)	18

Revision der MDR / IVDR

Kommentierung und Vorschläge aus der Sicht
der akademischen Forschung und
akademischen Verbundforschung



fit4translation

10.3	Datenschutz	18
11	Etablierung von einsehbaren Registern	19
11.1	Ist (aktuelle Regelung):	19
11.2	Problem.....	19
11.3	Vorschlag für eine Regelung	19
12	Transparenz und Feedback	19
13	Ansprechpartner für Rückfragen	20
14	Unterzeichnende Verbände	21

3 Nicht risikoangepasste Klassifizierung von Software (Anhang VIII Regel 11)

3.1 Ist (aktuelle Regelung):

Anhang VIII Regel 11 der MDR besagt, dass Software, die Informationen bereitstellt, die für diagnostische oder therapeutische Entscheidungen verwendet werden, grundsätzlich in Klasse IIa eingestuft wird. Wenn solche Entscheidungen eine potenziell schwerwiegende oder lebensbedrohliche Situation betreffen, kann die Einstufung in Klasse IIb oder sogar III erfolgen. Damit fallen sehr viele medizinische Softwareprodukte, insbesondere Clinical Decision Support Software (CDSS), automatisch in mittlere bis hohe Risikoklassen.

3.2 Problem

Die aktuelle Regelung berücksichtigt nicht ausreichend den Kontext der Nutzung sowie die spezifische Zweckbestimmung von Software. Dadurch werden Anwendungen, die nur mittelbar auf klinische Entscheidungen wirken (z. B. indem sie Empfehlungen geben, Informationen strukturieren oder visualisieren) oder deren Ergebnisse leicht durch den Anwender verifizierbar sind (z. B. im Falle KI-unterstützter Bildanalysesysteme), ähnlich streng reguliert wie Systeme, die direkt in den Behandlungsprozess eingreifen oder von deren Funktion das Überleben abhängt (z. B. Beatmungsgeräte).

Die Food and Drug Administration (FDA) in den USA differenziert hier stärker: Viele CDSS-Systeme (Clinical Decision Support Systeme) sind inzwischen ganz von der Regulierung ausgenommen, sofern der/die Anwender*in die Informationen unabhängig bewerten und eigene klinische Entscheidungen treffen kann. In der EU hingegen führt dieselbe Funktionalität häufig zu einer Klassifizierung als Hochrisikoprodukt. Diese Inkonsistenz ist schwer nachvollziehbar.

3.3 Vorschlag für eine Regelung:

Die Regel 11 in Anhang VIII der MDR sollte so angepasst werden, dass die Fehlerwahrscheinlichkeit (bei Bedienfehlern, Datenqualität) und die Wahrscheinlichkeit einen Fehler zu erkennen in die Risikoklassifizierung von Software und insbesondere CDSS aufgenommen werden. Einen ersten Ansatz bietet die Guidance on Qualification and Classification of Software der Medical Device Coordination Group (MDCG 2019-11 rev. 1), die in Anhang III die International Medical Device Regulators Forum (IMDRF) Risk Classification auf die Risikoklassen der MDR abbildet. Die IMDRF Risk Classification berücksichtigt die Relevanz der Informationen für Diagnose und Therapie, sowie die Kritikalität der Situation oder des Patientenstatus. Was die IMDRF Risk Classification nicht berücksichtigt, sind Kontext, die Datenherkunft und die spezifische Zweckbestimmung, die diese Faktoren des soziotechnischen Systems einbezieht. Es macht einen relevanten Unterschied, ob Daten einer gerade durch die Ärztin oder Arzt durchgeführten Behandlung in einer Notaufnahme / Ambulanz zusammengefasst und bewertet werden (alle Fakten den Nutzer*innen bekannt) oder ob durch die gleiche Software ein Arztbrief auf der Basis der Daten in KIS/KAS automatisch erstellt wird (Fakten häufig nicht alle Daten bekannt).



Beispiele in denen die Risikoklassifizierung anzupassen ist:

Beispiel 1: Eine Software, die bekannte und weithin akzeptierte Leitlinien zur Medikation herausgesucht und auf der Basis einer manuellen Eingabe der Patientendaten anzeigt, birgt ein geringes Risiko, da Ärzt*innen die eingegebenen Daten kennen, die Empfehlungen prüfen und die Informationen leicht auf Plausibilität bewerten können. In diesem Fall ist die Eintrittswahrscheinlichkeit gering, das Schadensausmaß ist theoretisch tödlich. Trotzdem ist das Gesamtrisiko gering.

Beispiel 2: Bei Software-Systemen, die auf der Grundlage von Fakten arbeiten, welche den Anwenderinnen und Anwendern sicher bekannt sind – beispielsweise weil sie diese Daten selbst eingegeben haben oder weil sie sich auf ein gerade geführtes und dokumentiertes Anamnesegespräch beziehen – können Fehler bei der Ergebnisgenerierung auftreten, etwa durch unzutreffende Diagnosevorschläge. Der entscheidende Unterschied ist jedoch, dass die zugrunde liegenden Informationen für die Ärztinnen und Ärzte transparent und nachvollziehbar sind. Dadurch können sie die Vollständigkeit und Plausibilität dieser Fakten prüfen und eventuelle Unstimmigkeiten erkennen. Das System hat eine hohe Entdeckungswahrscheinlichkeit. Damit ist das Risiko schwerwiegender Fehlinterpretationen als gering einzustufen. Werden von dem System Daten verwendet, deren Vollständigkeit und Plausibilität von den Nutzer*innen nur eingeschränkt verwertet werden können, zum Beispiel bei Systemen, die die Fakten für eine Entscheidung „zusammentragen“, ggf. sogar aus anderen Datenquellen erschließen müssen, ist die Entdeckungswahrscheinlichkeit geringer und daher das Risiko höher einzuschätzen. Verstärkt wird dies, wenn die Systeme proaktive durch Alerts oder Alarme auf kritische Situationen aufmerksam machen.

Beispiel 3: Systeme, die auf sogenannten „Black Box“-Algorithmen basieren, sind schwer nachvollziehbar, weil ihre Entscheidungsprozesse für die Anwenderinnen und Anwender nicht transparent sind. Dies führt dazu, dass Fehler oder fehlerhafte Ergebnisse nur mit geringer Wahrscheinlichkeit entdeckt werden können. Solche Systeme sind daher einer höheren Risikoklasse zuzuordnen. Systeme, deren Entscheidungsprozesse hingegen nachvollziehbar und erklärbar sind – etwa durch den Einsatz von erklärbarer Künstlicher Intelligenz (EAI) – ermöglichen es den Anwenderinnen und Anwendern, die automatisch erzeugten Ergebnisse zu überprüfen. Ein „Black Box“-System könnte eine Gewebeprobe automatisch einer bestimmten Tumorart zuordnen, ohne dass für die Ärztin oder den Arzt nachvollziehbar wird, auf welchen Mustern oder Kriterien diese Klassifizierung basiert. Demgegenüber kann ein System mit erklärbarer KI aufzeigen, welche Zellstrukturen, Färbungen oder Gewebemuster zur Entscheidung geführt haben. Dadurch kann die Pathologin oder der Pathologe diese Ergebnisse plausibilisieren und die eigene Diagnose mit einer Software / einem KI-System absichern. Bei solchen transparenten Systemen ist das Risiko geringer, weshalb sie einer niedrigeren Risikoklasse zugeordnet werden sollten.

Beispiel 4: Softwarelösungen, die Patient-Reported Outcome Measures (PROMs) – also direkt von Patientinnen und Patienten erhobene Angaben beispielsweise zu Schmerz, Fatigue oder Lebensqualität – automatisiert erfassen und auswerten, weisen unterschiedliche Risikoprofile auf. Wenn diese Daten direkt von den Patientinnen und Patienten eingegeben werden und Ärztinnen und Ärzte sie zeitnah im Rahmen einer Verlaufskontrolle prüfen und in den Gesamtkontext der Behandlung einordnen können, ist das Fehlerrisiko gering. In solchen Fällen ist eine Einstufung in eine niedrige Risikoklasse angemessen, gegebenenfalls kann sogar auf eine Klassifizierung als Medizinprodukt verzichtet werden. Ein deutlich höheres Risiko besteht jedoch, wenn PROMs automatisiert in klinische Entscheidungsunterstützungssysteme einfließen oder ärztliche Berichte generiert werden, ohne dass klar erkennbar ist, ob die Angaben vollständig und korrekt sind. In diesem Szenario steigt die Wahrscheinlichkeit fehlerhafter Ergebnisse, während die Möglichkeit der Fehlererkennung abnimmt. Eine höhere Risikoklasse ist hier daher gerechtfertigt.

In der aktuellen Regulierung ist erst die Risikoklassifizierung durchzuführen und erst im Risikomanagement danach werden die anderen Faktoren berücksichtigt. Diese haben dann keinen Einfluss mehr auf die Risikoklasse – und den darauf resultierenden regulatorischen Aufwand. Daher schlagen wir vor die Regel 11 von Anhang VIII der MDR und die IMDRF Risk Classification um folgende Aspekte zu erweitern und damit mehr Produkte den Klassen I oder IIa zuordnen.

- Einfluss auf und Relevanz für klinische Entscheidung
- Wahrscheinlichkeit Fehler zu erkennen
- Grad der Autonomie der Nutzer*innen bzw. deren Abhängigkeit vom System bei Entscheidungen.
- Qualifikation der Nutzer*innen, insb. Differenzierung der Anwendung durch Laien vs. Anwendung durch med. Fachpersonal
- Technische Komplexität: Ausnahmeregelung für einfache Score Berechnungen (z.B. Glasgow Coma Scale) geben. Diese sollten wie bei der FDA-Regulierung nicht als Medizinprodukt eingestuft werden oder Klasse I sein.

Damit würde eine feinere, realitätsnähere und innovationsfreundlichere Risikobewertung ermöglicht, die zugleich die Patientensicherheit wahrt.

4 Klinische Bewertung

4.1 Antizipierende klinische Bewertung und Stärkung der Überwachung im Feld

4.1.1 Ist (aktuelle Regelung)

Die MDR fordert eine umfassende klinische Bewertung bereits vor der Zulassung. Diese Bewertung muss den Nachweis erbringen, dass das Produkt die grundlegenden Sicherheits- und Leistungsanforderungen erfüllt. Fragen, die nicht in der klinische Bewertung vor der Inverkehrbringung beantwortet werden können, sind durch ein Vigilanzsystem und durch proaktive Post-Market Clinical Follow-up (PMCF) verpflichtend zu beantworten.



In der Logik wird bei der klinischen Bewertung und den Konformitätsbewertungsverfahren immer von deterministischen, sich nicht verändernden Produkten und Systemumgebungen ausgegangen. Für Software und KI-gestützte Anwendungen, werden daher in der Regel dieselbe Evidenzstrukturen erwartet wie für klassische Medizinprodukte.

4.1.2 Problem

Software und insbesondere Software mit KI-Komponenten sowie die Umgebung in der Software und KI-Systeme eingesetzt werden entsprechen jedoch häufig nicht diesen Kriterien. Anders als bei physischen Medizinprodukten ändern sich Softwareprodukte kontinuierlich:

- Regelmäßige Updates und Weiterentwicklungen können die Leistungsfähigkeit und das Risikoprofil erheblich verändern. Diese sind häufig allein aus IT-Security-Gründen erforderlich. Es ist für Entwickler und Hersteller ein erheblicher Mehraufwand dies für alle Versionen auch der abhängig verwendeten Software zu pflegen, daher sind häufig auch Releasewechsel erforderlich.
- KI-Algorithmen, insbesondere adaptive Modelle, können sich sogar im Betrieb dynamisch verändern oder bei kleinsten Änderungen der IT-Infrastruktur / Datenerfassung andere Ergebnisse ausgeben. Eine weitere Herausforderung ist, dass Ergebnisse bei generativer KI häufig nur eingeschränkt reproduzierbar sind.
- Eine vollständige klinische Bewertung ex ante ist unter diesen Bedingungen kaum möglich oder führt zu überproportional hohem Aufwand, ohne dass die Patientensicherheit dadurch real gesteigert wird. Die MDCG 2020-1 bietet hier bereits Möglichkeiten eines Konformitätsnachweis ohne klinische Prüfung, dies findet aber häufig keine Anwendung in der Praxis. Wünschenswert wäre, dass Teile der Überwachung gezielt in dem Post-market clinical follow-up (PMCF) durchgeführt werden können.

Die Folge: Hersteller sind gezwungen, umfangreiche klinische Bewertungen schon vor Markteintritt vorzulegen, obwohl ein Großteil der relevanten Erkenntnisse erst im realen Einsatz entstehen kann. Gleichzeitig werden PMCF und Vigilanz im aktuellen Regulierungsrahmen zu wenig genutzt, um diese Dynamik adäquat abzubilden. Dies gilt auch für Eigenherstellungen.

4.1.3 Vorschlag für eine Regelung

Die MDR sollte den in der Guidance on Clinical Evaluation der MDCG 2020-1 beschriebenen Ansatz weiter stärken. Ziel ist es, die Maßnahmen vor der Inverkehrbringung bzw. Inbetriebnahme besser mit den Anforderungen an die Überwachung im Feld abzustimmen:

- 1) Ex-ante Inverkehrbringung / Inbetriebnahme: Darstellung der grundsätzlichen Leistungsfähigkeit (insb. Usability, Relevanz, bei Software auch Genauigkeit). Bei Software, digitalen Biomarkern und KI sollten Leistungsprüfungen auch als in-silicio-Studien durchgeführt werden können. Für den Vergleich mit dem Stand der Technik auf der Basis von Literatur sollten klarere Kriterien veröffentlicht werden.
- 2) Ex-post Inverkehrbringung / Inbetriebnahme: PMCF und Vigilanz sollten für Software – insbesondere für KI – deutlich gestärkt und verpflichtender ausgestaltet werden. Dazu gehören:

- a. kontinuierliche Überwachung der Leistungsfähigkeit im Feld (auch Bewertung von Stichproben)
 - b. Nutzung von Real-World Data (z. B. Register, kuratierte Routinedaten - z.B. aus den Datenintegrationszentren der Medizininformatik-Initiative oder dem EHDS), für Training und kontinuierliche Validierung
 - c. klare Regeln für die Dokumentation und Kommunikation von Veränderungen (Software-Updates, Algorithmus-Anpassungen),
- 3) Strukturelle Verbesserung von Feedbackmechanismen, um erkannte Risiken schnell zu adressieren. Dies könnte z.B. durch die Integration von Meldungen aus der Medical Device Software heraus erfolgen bzw. die Software könnte Stichproben zur Qualitätssicherung vorsehen.

Dynamisches Genehmigungsmodell: Für Software mit regelmäßigem Update- oder Lernzyklus sollte ein regulatorischer Rahmen eingeführt werden, der Updates zulässt, solange die Grundfunktion unverändert bleibt und ein robustes PMCF/Vigilanz-System nachweislich etabliert ist.

Damit würde die klinische Bewertung besser an die Lebensrealität von Software und KI angepasst: weniger Überforderung durch unrealistische Anforderungen ex ante und mehr Patientensicherheit durch systematische Nachbeobachtung ex post. Ansonsten besteht das Risiko für die Patientensicherheit, da Verbesserungen zu spät in der Routineversorgung ankommen würden.

Daher unterstützen wir ausdrücklich die Bestrebungen der IMDRF zur Entwicklung der „Guidance on pre-determined change control plans (PCCP)“. Bei der Überarbeitung der MDR / IVDR sollte eine Konformität hergestellt werden.

4.2 Anpassung von Studiendesign Voraussetzungen an Risiko-Klassifizierung bei klinischen Studien und Leistungsprüfungen von Software, digitalen Biomarkern und KI-Systemen

4.2.1 Ist-Zustand (aktuelle Regelung)

Die MDR schreibt in Artikel 61 sowie Anhang XIV die Durchführung einer klinischen Bewertung und – sofern für die Bewertung erforderlich – klinische Prüfungen vor. Dabei müssen klinische Daten herangezogen werden, um Sicherheit, Leistungsfähigkeit und klinischen Nutzen eines Medizinprodukts nachzuweisen. In der Praxis verlangen viele Notified Bodies nicht nur bei Hochrisiko-Produkten (Klasse III, implantierbar) prospektive Studiendaten, auch wenn dies in der MDR nicht explizit verlangt wird. Außerdem bleibt unklar, welche Studientypen für KI-Systeme der Risikoklassen I und II als „angemessen“ gelten. Der AI-Act ergänzt die MDR zwar durch Anforderungen an Risikomanagement und Monitoring, definiert aber ebenfalls keine klaren Standards für klinische Validierung.

4.2.2 Problem:

Die fehlende Klarheit darüber, welche Studiendesigns für die Validierung medizinischer KI-Systeme geeignet sind, führt zu erheblichen Risiken. Bei medizinischen KI-Systemen kommt im Vergleich zu

herkömmlichen Medizinprodukten hinzu, dass retrospektive Daten für eine Validierung herangezogen werden können. Retrospektive Studien können helfen den Arbeitsaufwand zur Validierung gering zu halten, erfassen jedoch nicht die tatsächliche Wirkung im medizinischen Alltag, etwa die Interaktion mit klinischen Workflows oder die Auswirkungen auf Patient*innen. Für Niedrig-Risiko-Systeme oder die erneute Prüfung nach Software-Updates von zugelassener Software können retrospektive Analysen ausreichend sein. Allerdings kann eine fehlende prospektive Validierung bei KI-Systemen mit höherem Risiko (Klasse II und III) zu einer gefährdeten Patientensicherheit führen, da belastbare Daten zu Nutzen und Risiken im realen Einsatz fehlen. Für Medizinprodukte ab Risiko-Klasse II bleibt unklar nach welchen Maßstäben die vorliegende Evidenz von den Notified Bodies beurteilt werden – eine zugängliche Einsicht der Bewertungsgrundlage nach Zulassung gibt es nicht. Diese Unklarheit führt zu Herausforderungen für Softwareentwickelnde Institutionen, um eine angemessene Validierungsstrategie vorzubereiten. Darüber hinaus bleibt für klinische Anwender*innen unklar, auf welcher Qualität der Evidenz die Sicherheit eines Systems tatsächlich beruht.

4.2.3 Vorschlag für eine Regelung

Die klinische Validierung medizinischer KI-Systeme sollte klar an ihre Risikoklassifizierung gekoppelt werden, um MDR/IVDR und AI-Act zu harmonisieren. Für Hochrisiko-Systeme (MDR Klasse III sowie bestimmte Klasse IIb) sind verpflichtend prospektive klinische Studien erforderlich, ergänzt durch retrospektive Daten, während für Systeme mit mittlerem Risiko (vor allem Klasse IIa und Teile von Klasse IIb) eine Kombination aus retrospektiven und prospektiven Beobachtungsdaten ausreichen kann. Niedrigrisiko-Systeme (Klasse I) können durch retrospektive Studien oder Benchmark-Datensätze validiert werden, sofern Sicherheit und Leistungsfähigkeit nachgewiesen sind. Für Software-Updates und neue Versionen sollen retrospektive Benchmark-Validierungen ausdrücklich anerkannt werden, wobei prospektive Zusatzstudien nur bei wesentlichen Funktionsänderungen nötig sind. Diese risikoadaptierte Grundlage schafft Patientensicherheit, regulatorische Klarheit, Kohärenz zwischen MDR/IVDR und AI-Act und fördert gleichzeitig Innovation sowie verlässliche Anwendbarkeit im klinischen Alltag.

5 Klinische Prüfungen – Einwilligung bei Cluster Randomisierten Studien

5.1 Ist (aktuelle Regelung)

Artikel 63 MDR schreibt vor, dass eine klinische Prüfung mit Medizinprodukten grundsätzlich nur durchgeführt werden darf, wenn die betroffenen Patient*innen vorab ihre informierte Einwilligung gegeben haben. Diese Regelung entspricht dem Grundprinzip der klinischen Forschungsethik, eine individuelle Zustimmung einzuholen.

5.2 Problem

Für Medical Device Software und insbesondere Clinical Decision Support Systems (CDSS) ist diese Anforderung in vielen Studiendesigns praktisch nicht umsetzbar: CDSS wirken häufig nur über die Beratung von Ärztinnen oder über automatisierte Hinweise in IT-Systemen. Die Patientinnen

interagieren nicht direkt mit der Software. Solche Software wird typischerweise auf der Ebene ganzer Einrichtungen oder Abteilungen eingeführt (z. B. Integration in Informationssysteme im Krankenhaus). Eine individuelle Auswahl, für welche Patient*innen die Software von dem Klinikpersonal genutzt werden darf und für welche nicht, ist technisch nicht realisierbar. Dies kann bei Systemen, die proaktiv Hinweise, Alerts oder Alarmer ausgeben gefährlich werden, da sich die Anwender auf die Ausgaben des Systems verlassen. Weiter haben solche Systeme auch weitergehende Einflüsse auf das soziotechnische System und damit die Versorgung, so dass man diese Systeme nur gesamt für eine organisatorische Einheit einführen kann bzw. sollte. In diesem Fall erfolgt die Evaluation der Systeme in der Regel über Studien mit einer Cluster-Randomisierung, bzw. einem Stepped-Wedge-Design. Eine individuelle Einwilligung wäre hier nur mit unverhältnismäßigem Aufwand oder faktisch gar nicht möglich.

Folge: Ohne eine Ausnahmeregelung von der individuellen Einwilligungspflicht werden solche Studien – die für die realistische Evaluation digitaler Systeme besonders geeignet sind – praktisch verhindert.

5.3 Vorgeschlagene Regelung

Art. 63 MDR sollte um eine spezifische Ausnahme für Medizinprodukte, die nur in cluster-randomisierte Studien sinnvoll untersucht werden können, erweitert werden:

- **Ausnahme von der individuellen Einwilligung:** Wenn eine Software ausschließlich indirekt auf Patient*innen wirkt (z. B. durch ärztliche Entscheidungsunterstützung), kann auf die individuelle Einwilligung verzichtet werden. Stattdessen erfolgt eine Information der Patientinnen über das Studiendesign sowie die Möglichkeit zum Widerspruch („Opt-Out“).
- **Pseudonymisierte Datenverarbeitung:** Für die wissenschaftliche Auswertung darf eine Verarbeitung pseudonymer Daten erfolgen, sofern angemessene Datenschutz- und Sicherheitsmaßnahmen bestehen. (Bsp. § 6 Absatz 3 Satz 4 Gesundheitsdatennutzungsgesetz GDNG)
- **Öffnungsklausel für nationale Regelungen:** Mitgliedstaaten können über eine Öffnungsklausel weitergehende Anforderungen oder Schutzmechanismen festlegen (z. B. wie im deutschen GDNG vorgesehen).

Positive Wirkung: Mit einer Ausnahmeregelung für Clusterrandomisierte Studien auf der Basis von Routinedaten werden cluster-randomisierte und Stepped-wedge-Studien zu Software und CDSS überhaupt erst praktikabel. Dies ermöglicht valide Evidenz zur Wirksamkeit und Sicherheit in der realen Versorgungspraxis, ohne Patient*innen unnötig zu belasten oder Studien methodisch zu blockieren. Weiter ist dies auch eine gute Lösung für die kontinuierliche Überwachung (Monitoring) von in Betrieb befindlichen KI-Systemen.

6 Eigenherstellung nach Art. 5(5) MDR

6.1 Gewinnung von „Erfahrungen“ zum Einsatz von Eigenherstellungen

6.1.1 Ist (aktuelle Regelung)

Artikel 5 Absatz 5 (h) MDR verpflichtet Gesundheitseinrichtungen, die aus der klinischen Verwendung von Eigenherstellungen gewonnenen Erfahrungen zu begutachten und erforderliche

Korrekturmaßnahmen zu ergreifen. Damit soll die kontinuierliche Sicherheit und Leistungsfähigkeit von Eigenherstellungen gewährleistet werden.

6.1.2 Problem

Bei Software und insbesondere bei KI-Systemen ist eine kontinuierliche Überwachung der Leistung und Sicherheit erforderlich:

- **Soziotechnische Abhängigkeit:** Die Leistung und Sicherheit von Software und KI hängt stark vom Anwendungskontext und der technischen Umgebung ab (z. B. Einbindung in IT-Landschaft, Datenqualität inkl. Interoperabilität, Prozesse und Kompetenzen der Anwender)
- **Kurze Lebenszyklen:** Software und KI-Module werden in sehr kurzen Abständen aktualisiert oder verändert, auch um die Cybersecurity sicherzustellen. Damit verändert sich ggf. auch das Verhalten.

Dies erfordert eine verstärkte, systematische Überwachung der Anwendung der Systeme im Ziel. Dabei besteht die Gefahr einer regulatorischen Einstufung der Datenerhebung als klinische Prüfung im Sinne von Artikel 2 Nr. 45 MDR. Dies würde zum Verlust der Privilegien der Eigenherstellung und Einstufung als Prüfprodukt (Artikel 2 Nr. 46 MDR) führen. Daraus resultieren zusätzliche Anforderungen und damit ein erheblicher bürokratischer Aufwand – auch dann, wenn es nur um eine routinemäßige Überwachung der Sicherheit im klinischen Alltag geht.

Das Risiko: Sinnvolle und für die Patientensicherheit notwendige Datenerhebungen bzw. Studien zur Überwachung der Eigenherstellungen im Betrieb werden erschwert oder ganz verhindert. Damit schwächt die Konstellation das eigentliche Schutzziel.

6.1.3 Vorschlag für eine Regelung:

Für die Evaluation von bereits implementierten Eigenherstellungen sollte eine klare Ausnahme analog zu Artikel 74 MDR eingeführt werden: Studien und Untersuchungen zur Leistungsfähigkeit und Sicherheit im Routinebetrieb sollten nicht als klinische Prüfung unter der Regelung von Artikel 62 ff oder Artikel 82 MDR gelten, außer wenn in der Studie invasive oder belastende studienbedingte Maßnahmen durchgeführt werden (Regelung analog zu Artikel 74 MDR). Bei einer Publikation der Ergebnisse ist vor dem Beginn der Datensammlung die Studie in einer nach nationalen (Landes-) Recht gebildeten Ethik-Kommission zu beraten und vor Beginn der Datensammlung in einem WHO-Primärregister zu registrieren.

Für Eigenherstellungen sollte klargestellt werden, dass systematische Untersuchungen bzw. Überwachungsmaßnahmen im realen Anwendungskontext (z. B. Monitoring, Audits, Nutzerfeedback, technische Leistungsüberwachung) als Teil der Überwachungs- und Qualitätssicherungsprozesse zu behandeln sind und nicht als klinische Prüfung – unabhängig ob die Daten publiziert werden sollen. Hier sollten die Empfehlungen der Guidance on the health institution exemption under Article 5 (5) der Medical Device Coordination Group (MDCG-Guidance 2023-01) weiter ausgeführt werden.

Damit wird ermöglicht, dass Gesundheitseinrichtungen regelmäßig systematische Studien und Evaluationen zur Patientensicherheit und Wirksamkeit ihrer Eigenherstellungen durchführen können, ohne durch die Hürden einer formalen klinischen Prüfung blockiert zu werden. Dies fördert eine

gelebte Sicherheitskultur und trägt zur kontinuierlichen Verbesserung von Software und KI im Versorgungsalltag bei.

6.2 Einheitliche Regelung für Hochschulmedizin im Kooperationsmodell / Klinikverbund, Praxisnetzwerke und öffentlich geförderte Verbundforschungsprojekte

6.2.1 Ist (aktuelle Regelung)

Nach Artikel 5 Absatz 5 (a) MDR dürfen Eigenherstellungen von Gesundheitseinrichtungen nicht an eine andere rechtlich eigenständige Einrichtung abgegeben werden. Die Guidance on the health institution exemption under Article 5 (5) der Medical Device Coordination Group (MDCG-Guidance 2023-01) bestätigt diese strikte Auslegung. Ziel der Regelung ist es, die Ausnahme für Eigenherstellungen eng zu halten und eine Umgehung der regulären Marktzulassung zu verhindern.

6.2.2 Problem

In Deutschland ist die Universitätsmedizin in Integrationsmodellen und Kooperationsmodellen zwischen den Unikliniken organisiert. Während die Standorte mit einem Integrationsmodell die Kompetenzen ihrer (theoretischen) Institute für die Entwicklung von Eigenherstellungen nutzen kann, ist dieser Weg bei Kooperationsmodellen zunächst verschlossen. Ebenfalls ist die in Verbundprojekten übliche, institutsübergreifende gemeinsame Entwicklung nicht möglich.

Auch in der Versorgung ergeben sich erhebliche Probleme mit der Klausel für die gemischt stationär und ambulant tätigen Versorgungseinrichtungen. Bei einem Institut (Labor, Pathologie), das sich die Räumlichkeiten und Einrichtung mit einem MVZ teilt, müssen formal die Vorgaben zur Eigenherstellung für jede rechtlich eigenständige Institution einzeln umgesetzt werden, obwohl beide Institutionen dieselben Ressourcen nutzen.

Die Folge: Strukturelle Unterschiede in der Rechtsform führen zu einer Ungleichbehandlung von Standorten und behindern die Verbesserung der Patientenversorgung durch Eigenherstellungen.

6.2.3 Vorschlag für eine Regelung

Die Potentiale der Kompetenzen der Hochschulen und einer gemeinsamen Eigenherstellung mehrerer Gesundheitseinrichtungen sollten für eine optimale Versorgung und Patientensicherheit genutzt und daher ebenfalls privilegiert werden.

Daher sollte es für Universitätskliniken die Möglichkeit geben, mit der dazugehörigen Universität sowie assoziierten akademischen Lehrkrankenhäusern Eigenherstellungen entwickeln und betreiben zu können. Für diese Konstellationen, sowie eine gemeinsame Entwicklung und Nutzung von öffentlichen Gesundheitseinrichtungen und Hochschulen / Forschungseinrichtungen sowie die Eigenherstellung innerhalb eines Standortes, der von mehreren Einrichtungen genutzt wird, sollte es die Möglichkeit einer Eigenherstellung in einer gemeinsamen Verantwortung geben, in dem die behandelnde Klinik Einrichtung gegenüber dem Patienten haftet und die Verantwortlichkeiten im Binnenverhältnis vertraglich zu regeln sind.

6.3 Abgabe als Open Source Software

6.3.1 Ist (aktuelle Regelung)

Die Artikel 5 Absatz 5 MDR fordert, dass eine Eigenherstellung nicht an eine andere rechtlich eigenständige Einrichtung abgegeben werden darf.

6.3.2 Problem

Streng genommen gilt dies auch für die Veröffentlichung von Open-Source-Software, auch wenn diese nicht unmittelbar auf dem Markt bereitgestellt wird oder in dieser Form direkt als Medizinprodukt nutzbar ist. Gleichwohl ist es wünschenswert, dass selbst hergestellte Software im Sinne von Open Science entsprechend veröffentlicht werden kann. Dies ist insbesondere deshalb wichtig, weil viele Open-Source-Bibliotheken unter einer Copyleft-Lizenz veröffentlicht werden, die eine Veröffentlichung der Software als Open Source Software (OSS) vorschreibt.

6.3.3 Vorschlag für eine Regelung

Es sollte klargestellt werden, dass Eigenherstellungen als Open Source Software verfügbar gemacht werden dürfen, damit die Innovationen und Ideen aus den Eigenherstellungen auch von anderen Einrichtungen genutzt werden können. Eine entsprechende Auslegung von Artikel 5 Absatz 5 MDR ist möglich, birgt aber Rechtsunsicherheiten. Daher ist hier eine Klarstellung in der MDR oder eine Klarstellung über eine Guideline der MDCG (z.B. der MDCG 2023-1) wünschenswert.

7 Eigenherstellungen nach IVDR

Die Unterzeichner schließen sich der Stellungnahme der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) im Rahmen der gezielten Bewertung der Verordnung 2017/746 über In-vitro-Diagnostika gemäß Artikel 111 und Verordnung (EU) 2024/1860 vp, 21. Februar 2025 an (Stellungnahme in Anlage A). Insbesondere sehen wir einen Handlungsbedarf für eine Neureglung von Artikel 5 Absatz 5 IVDR, der sowohl die Krankenversorgung als auch die Innovationsfähigkeit der Universitätsmedizin gefährdet.

8 Öffentlich zugängliche Informationen über medizinische KI-Systeme

8.1 Ist (aktuelle Regelung)

Unter der MDR/IVDR sind wesentliche technische Informationen (z. B. vollständige Zweckbestimmung, Trainingsdaten, Validierungsdaten) nur in der Technischen Dokumentation für Notified Bodies und Behörden zugänglich. Öffentlich verfügbar sind in der Regel lediglich eine gekürzte Zweckbestimmung in der Gebrauchsanweisung sowie bei höheren Risikoklassen eine „*Summary of Safety and Clinical Performance (SSCP)*“. Der AI-Act schreibt zwar umfangreiche Dokumentationspflichten für Hoch-Risiko-Systeme vor, richtet diese jedoch ebenfalls primär an Behörden. Transparenzpflichten für General Purpose AI (GPAI) sind stärker ausgeprägt, insbesondere im Hinblick auf Trainingsdaten. Für medizinische Hoch-Risiko-Systeme bestehen vergleichbare Anforderungen bislang nicht. Für Anwender*innen und Forscher*innen bedeutet dies, dass sie derzeit nur sehr eingeschränkt Einblick in Sicherheit, Trainingsdaten und Limitationen solcher Systeme erhalten (siehe unsere Publikation: <https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2024.1267290/full>).



8.2 Problem:

Unsere Analyse zeigt, dass die öffentlich zugänglichen Informationen von KI-Systemen in der Risiko-Klasse IIb derzeit nicht ausreichen, um die Sicherheit und Eignung von KI-Systemen im klinischen Alltag zuverlässig zu beurteilen. Zwar existiert die Zweckbestimmung als zentrales Steuerungsinstrument, diese ist jedoch häufig nicht öffentlich einsehbar oder so allgemein formuliert, dass eine verlässliche Einschätzung des tatsächlichen Einsatzkontexts nicht möglich ist. Daraus ergeben sich mehrere Risiken: Anwender*innen können vor dem Einsatz nur schwer oder gar nicht einschätzen, ob ein KI-System für ihre spezifische Patientenpopulation geeignet ist, beispielsweise für Patient*innen mit seltenen Vorerkrankungen, speziellen Symptomen oder aus bestimmten demographischen Gruppen. Zudem fehlen oft Angaben zu Validierungsgrenzen, sodass unklar bleibt, in welchen klinischen Umgebungen, mit welchen bildgebenden Geräten oder bei welchen Subgruppen die Software getestet wurde oder nicht getestet wurde. Auch die externe Überwachung durch Forschungseinrichtungen ist eingeschränkt, da diese keinen Zugang zu den notwendigen Daten haben und Systeme daher nicht unabhängig überwachen oder evaluieren können. Darüber hinaus besteht eine regulatorische Inkohärenz: Während MDR/IVDR vor allem auf nicht-öffentliche technische Dokumentation setzen, adressiert der AI-Act Transparenzpflichten primär gegenüber Behörden und für General Purpose AI. Ohne Anpassung entstehen dadurch Widersprüche und Transparenzlücken für medizinische Hoch-Risiko-Systeme.

8.3 Vorschlag für eine Regelung

Wir schlagen vor, dass die vollständig ausformulierte Zweckbestimmung medizinischer KI-Systeme verpflichtend veröffentlicht wird. Diese Zweckbestimmung sollte eine klare Definition des Einsatzkontextes enthalten, einschließlich der relevanten Patientengruppen, Indikationen, klinischen Settings und Validierungsgrenzen. Sie muss öffentlich zugänglich sein, zum Beispiel über die Gebrauchsanweisung (IFU), EUDAMED oder Produktdatenbanken. Darüber hinaus sollten Hersteller verpflichtet werden, die verwendeten Datensätze für Training und Validierung ihrer KI-Systeme öffentlich zu beschreiben. Die Dokumentation kann sich an etablierten Frameworks wie TRIPOD+AI, STARD-AI, CONSORT-AI oder MINIMAR orientieren und sollte Informationen zu Datenquellen und -arten, zur Abdeckung verschiedener demographischer Gruppen, zu getesteten klinischen Szenarien und Geräten sowie zu bekannten Limitationen enthalten. Diese Informationen müssen sowohl Anwender*innen als auch Forschungseinrichtungen zugänglich gemacht werden, um ein kontinuierliches klinisches Monitoring und unabhängige Forschung zu ermöglichen.

9 Open Source Software und SOUP

9.1 Ist (Aktuelle Regelung)

Nach MDR dürfen Hersteller nur Software einsetzen, deren Entwicklung und Qualität sie belegen können. Die Norm IEC 62304:2006 mit dem Amendment IEC 62304:2006/Amd 1:2015, bzw. der DIN EN 62304:2016-10 sieht den Umgang mit „Software of Unknown Provenance“ (SOUP) ausdrücklich vor – also Software, deren Entwicklungsprozess nicht vollständig dokumentiert ist (z. B. Open Source Libraries). Allerdings ist die DIN EN 62304:2016-10 bis heute nicht mit der MDR harmonisiert. Daraus resultieren viele Unsicherheiten bei der Einbindung von Open Source Software (OSS) oder SOUP-Komponenten in Medizinprodukten, Prüfprodukten oder Eigenherstellungen.

9.2 Problem

Die aktuelle Rechtslage entspricht nicht der Realität der Softwareentwicklung: OSS ist Standard. Nahezu alle modernen Softwareprodukte – insbesondere im Bereich KI und Datenanalyse – basieren auf Open-Source-Bibliotheken und -Frameworks (z. B. TensorFlow, PyTorch, NumPy). Ohne klare Regelung zur Verwendung von SOUP wäre praktisch jede Softwareentwicklung für Medizinprodukte unmöglich oder müsste proprietäre Eigenentwicklungen aller Basiskomponenten verlangen. Daraus resultiert eine regulatorische Unsicherheit: Entwickler*innen von Medizinsoftware, Prüfprodukten oder Eigenherstellungen bewegen sich im Graubereich, da es keine eindeutige rechtliche Grundlage für den Einsatz von SOUP/OSS gibt. Ein pauschales Verbot führt nicht zu mehr Sicherheit, sondern verhindert ein strukturiertes Risikomanagement und Monitoring, das SOUP/OSS-Komponenten tatsächlich sicher nutzbar machen könnte.

Vor allem zeigt die Erfahrung, dass bei kommerzieller Software zwar Verantwortlichkeiten vertraglich besser geregelt sein mögen, dies aber keine Garantie für einen höheren Sicherheits- und Transparenzlevel ist.

9.3 Vorgeschlagene Regelung

Die MDR sollte den Einsatz von SOUP / OSS ausdrücklich zulassen bzw. sollte eine Guideline der MDCG den Einsatz von SOUP / OSS erlauben, wenn für den Software Lifecycle des Gesamtsystems die in den Standards DIN EN IEC 62304:2016-10, DIN EN ISO 14971:2019-11, DIN EN IEC 81001, DIN EN IEC 81001-5-1:2023-07 beschriebenen Anforderungen erfüllt sind:

- 1) **Zulässigkeit von SOUP:** SOUP / OSS darf eingesetzt werden, wenn Risiken im Rahmen des Software-Entwicklungsprozesses systematisch identifiziert und kontrolliert werden. (entsprechend in der ISO 14971)
- 2) **Risikomanagement:** Für jede SOUP-Komponente ist eine Bewertung hinsichtlich Funktion, Kritikalität, Sicherheitsrelevanz und möglicher Schwachstellen durchzuführen.
- 3) **Kontinuierliches Monitoring:** Hersteller müssen ein kontinuierliches Schwachstellen-Monitoring etablieren (z. B. im Rahmen von Continuous Integration/Continuous Deployment-Pipelines), um bekannte Sicherheitslücken zeitnah zu erkennen und Updates einzuspielen.
- 4) **Dokumentation:** Anforderungen an Dokumentation sollten sich auf Nachvollziehbarkeit und Transparenz (Versionierung, Patch-Status, Sicherheitsbewertung, Reaktion auf bekannt

gewordene Schwachstellen und Sicherheitslücken) konzentrieren – nicht auf eine vollständige Entwicklungsdokumentation der Fremdsoftware.

Damit wird die rechtliche Sicherheit geschaffen, SOUP/OSS verantwortungsvoll in Medizinprodukten einzusetzen. Dies stärkt Innovationsfähigkeit, ermöglicht realistische KI- und Softwareentwicklung und erhöht die Sicherheit, da Schwachstellen systematisch überwacht und geschlossen werden. Alternativ wäre eine erneute Harmonisierung der ISO 62304 anzustreben.

10 Harmonisierung der MDR / IVDR mit anderen Richtlinien und Verordnungen

10.1 AI-Act

Insgesamt ist ein Beschleunigen der Anpassungen der verschiedenen Rechtsakte der EU und der untergesetzlichen Regelungen anzustreben. (Gruppe Interblade). Dies umfasst unter anderem:

- Artikel 11 AI-Act (Technische Dokumentation): Die Anforderungen sind sehr detailliert und umfangreich. Eine Harmonisierung ist dringend notwendig, damit technische Dokumentation unter MDR/IVDR und AI-Act nicht mehrfach und widersprüchlich geführt werden muss.
- Artikel 10 AI-Act (Daten und Daten Governance): Der AI-Act regelt Datenqualität, Bias etc. MDR/IVDR enthalten Anforderungen an klinische Daten, aber nicht in allen Fällen so spezifisch wie der AI-Act. Unterschiede können zu Inkonsistenzen führen.
- Artikel 12 AI-Act (Record-keeping / Nachverfolgbarkeit): MDR/IVDR haben Anforderungen an die technische Dokumentation und die Rückverfolgbarkeit, aber der AI-Act verlangt möglicherweise zusätzliche Aufzeichnungen (z. B. Logbücher über Entscheidungen, automatische Änderungen).
- Artikel 14 AI-Act (Human Oversight): MDR/IVDR verlangen, dass Geräte sicher im Gebrauch sind und Bedienung durch qualifiziertes Personal möglich ist, der AI-Act kann zusätzliche Anforderungen an Monitoring der Entscheidungen und Transparenz gegenüber den Nutzer*innen verlangen.
- Artikel 65–67 AI-Act (Post-Market Monitoring / Überwachung): MDR/IVDR enthalten Vigilanz und PMS, aber nach dem AI-Act können zusätzliche/andere Meldepflichten, Datenerhebung, Berichte etc. gefordert werden. Eine Harmonisierung hier hilft Redundanz zu vermeiden.
- Artikel 2 MDR Anhang VIII (Regel 11): Die MDR Anhang VIII Regel 11 für Software-Risikoklasse ist maßgeblich auch für AI-Systeme. Hier sollte Klarheit bestehen, wie das mit der Risiko-Kategorisierung unter AI-Act (Artikel 6) zusammenfällt. (siehe auch Abschnitt 2).
- MDR / IVDR technische Dokumentation Artikel (z. B. MDR Artikel 10-11 etc.): Diese Artikel und die Vorgaben für klinische Bewertung / Leistung / Sicherheit etc. müssen harmonisiert mit den Dokumentationspflichten im AI-Act sein, damit Hersteller nicht zwei unterschiedliche Dokumentationen pflegen müssen.

10.2 Cybersecurity (NIS2, KRITIS)

10.2.1 Sicherheit durch schnelle Reaktionen (Patches) versus Konformitätsverfahren

Die Sicherheit von Medizinprodukten, insbesondere im Bereich Cybersecurity, erfordert schnelle Updates und ein gut organisiertes Releasemanagement. Dabei dürfen die Prozesse weder die Entwickler und Hersteller noch die Betreiber und Anwender überfordern. Konformitätsbewertungsverfahren nach MDR/IVDR dauern in der Praxis deutlich länger, insbesondere wenn zuvor klinische Daten für die Bewertung gesammelt werden müssen. Gleichzeitig ist es nicht praktikabel, eine große Zahl von Versionsständen parallel zu pflegen oder Patches für viele Versionen zeitnah zu erstellen und zu testen. Eine sichere Nutzung von Medizinprodukten ist daher nur durch kontinuierliche Updates und ein laufendes Monitoring im Betrieb möglich. Dieses Monitoring muss sicherstellen, dass neu auftretende Fehler, Schwachstellen, Sicherheitslücken oder sonstige Vorkommnisse zeitnah erkannt und qualifiziert bewertet werden. Dieses Vorgehen unterstützt das Konzept der antizipierenden klinischen Bewertung und der Stärkung der Überwachung im Feld, wie sie in Abschnitt 3 beschrieben ist.

10.2.2 Weitere erforderliche Harmonisierung / Vermeidung von Doppeldokumentation

MDR/ IVDR und NIS2-Richtlinie / KRITIS stellen ähnliche, aber nicht identische Anforderungen an die Dokumentation. Es sollte angestrebt werden, dass die folgende Dokumentation in einer gemeinsamen Dokumentation erfolgen kann:

- Zentrale Dokumentation: Gemeinsame Vorlagen für die Meldungen von sicherheitsrelevanten Vorfällen an Betreiber, Behörden und Hersteller (Interoperabilität des Meldewesens), Patchmanagement, Risikomanagement, die sowohl MDR/IVDR als auch die NIS2-Richtlinie/KRITIS erfüllen.
- Abgestimmte Fristen: Harmonisierung von Reaktionszeiten auf Sicherheitsvorfälle, um Konflikte zwischen MDR, DSGVO, NIS2-Richtlinie und anderen Meldepflichten zu vermeiden.
- Klare Rollenaufteilung: Definition, welche Aufgaben Hersteller, Betreiber, Anwender übernehmen, ggf. durch Verträge oder gesetzliche Klarstellungen.
- Integration von Normen: MDR/IVDR-konforme Risikomanagement- und Software-dokumentation um NIS2- Anforderungen ergänzen, statt separat zu führen.
- Audit-Harmonisierung: Gemeinsame Prüfkriterien / Audit-Dokumentationen für verschiedene Zwecke (insb. bei Audits zu MDR/IVDR, Informationssicherheit / KRITIS, Datenschutz) um Doppelprüfungen zu vermeiden.

10.3 Datenschutz

Die Nutzung von personenbezogenen (Patient*innen) Daten für clusterrandomisierte Studien sowie die kontinuierliche Überwachung im Feld sollte auf die DSGVO und ggf. nationale Regelungen wie das GDNG oder die EHDS-Verordnung auf Europäischer Ebene verweisen. Durch MDR und IVDR sollten keine zusätzlichen Hürden aufgebaut werden (z.B. durch eine Einwilligungspflicht für die Datenverarbeitung), die Maßnahmen zur Verbesserung der Sicherheit von Medizinprodukten, Prüfprodukten und / oder Eigenherstellungen einschränkt.

11 Etablierung von einsehbaren Registern

11.1 Ist (aktuelle Regelung):

Die MDR sieht mit EUDAMED (European Database on Medical Devices) eine zentrale europäische Datenbank vor, die Informationen zu Medizinprodukten, Herstellern, Zertifikaten und Vigilanzmeldungen sammelt. Ziel ist es, Transparenz über zugelassene Produkte und deren regulatorischen Status zu schaffen. Der AI-Act schreibt ebenfalls Registrierungspflichten für Hochrisiko-KI-Systeme vor, die im EU-Markt eingesetzt werden. Diese Dokumentation bezieht sich allerdings primär auf die Zulassungsebene (Produktidentität, Hersteller, Risikoklasse) und nicht auf den konkreten Einsatz in Gesundheitseinrichtungen. Aktuell gibt es weder auf europäischer noch auf nationaler Ebene ein Register, das nachvollziehbar macht, wo ein KI-System im Gesundheitswesen tatsächlich implementiert ist, wodurch eine externe Überwachung verhindert wird.

11.2 Problem

Für Gesundheitseinrichtungen besteht derzeit eine erhebliche Intransparenz: Kliniken und Praxen können kaum nachvollziehen, in welchen Settings ein bestimmtes KI-System bereits erfolgreich eingesetzt wird, was fundierte Entscheidungen über Beschaffung und Implementierung erschwert. Gleichzeitig fehlt wissenschaftlichen Einrichtungen eine systematische Datengrundlage, um Wirksamkeit, Zuverlässigkeit und Kontextabhängigkeit von KI-Systemen in unterschiedlichen Versorgungsszenarien zu evaluieren. Damit klafft eine Lücke zwischen Regulierung und Praxis: Während MDR und AI Act Transparenz auf Produktebene gewährleisten, fehlt es an nachvollziehbarer Transparenz auf Anwendungsebene.

11.3 Vorschlag für eine Regelung

Wir schlagen die Etablierung eines nationalen, öffentlich zugänglichen Registers für medizinische KI-Systeme vor, idealerweise als Erweiterung von EUDAMED, das neben der Geräte-Registrierung spezifische Angaben zu Zweckbestimmung, Risikoklasse, Validierungsdaten und Limitationen enthält. Ergänzend dazu könnten nationale Register aufgebaut werden, die den konkreten Einsatz dokumentieren, etwa welche Systeme in welchen Krankenhäusern und Indikationen tatsächlich genutzt werden, um Kliniken und Behörden eine fundierte Entscheidungsgrundlage zu geben. Durch die Integration von Real-World-Evidenz-Strukturen ließe sich sicherstellen, dass auch Daten aus der Versorgungspraxis systematisch erfasst werden, etwa zu Performance, Bias oder Patientensicherheit. Darüber hinaus sollte das Register eine Feedback- und Monitoring-Funktion enthalten, die Rückmeldungen aus dem klinischen Alltag sammelt und so ein lernendes System etabliert. Auf diese Weise entstünde eine Infrastruktur, die Transparenz, Aufsicht und kontinuierliche Verbesserung verbindet und zugleich den Anforderungen von MDR und AI-Act gerecht wird.

12 Transparenz und Feedback

Die Veröffentlichung klarer Leitfäden durch die MDCG ist grundsätzlich hilfreich. Sie könnten jedoch noch deutlicher formulieren, was in den unterschiedlichen Szenarien als angemessen gilt – also

Revision der MDR / IVDR

Kommentierung und Vorschläge aus der Sicht
der akademischen Forschung und
akademischen Verbundforschung



fit4translation

praktisch ‚was gut genug ist‘. Dies betrifft insbesondere den Umfang der erforderlichen Dokumentation und die damit verbundene Bürokratie, sodass Hersteller und Betreiber klare Orientierung erhalten, ohne unnötig belastet zu werden. Hier wären auch Musterbeispiele für die Strukturierung von Dokumenten oder Templates, z.B. für den Plan für die Klinische Bewertung und den Bericht, oder den Plan für die Überwachung nach der Inverkehrbringung hilfreich.

Benannte Stellen und Behörden sollten verpflichtet werden, Rückmeldungen sowie öffentlich zugängliche Musterbescheide bereitzustellen. Dies würde Entwicklern, Herstellern und Forschern ermöglichen, besser zu verstehen, wie die Regelungen in der Praxis angewendet und geprüft werden.

13 Ansprechpartner für Rückfragen

TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e. V.
Charlottenstraße 42
10117 Berlin

Sebastian C. Semler, Geschäftsführer

Tel.: +49 (0)30 2200247-10

sebastian.semler@tmf-ev.de



fit4translation

14 Unterzeichnende Verbände



Die **TMF** – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. steht für Forschung, Vernetzung und Digitalisierung in der Medizin. 1999 gegründet, ist sie die Dachorganisation zur Digitalisierung in der medizinischen Verbundforschung in Deutschland, im Rahmen derer Spitzenforscherinnen und -forscher Wissen austauschen, gemeinsam Ideen und Konzepte entwickeln und so die Zukunft der medizinischen Forschung im digitalen Zeitalter gestalten.



fit4translation

und Usability-Labor etabliert, damit Produkte praxisnah geprüft werden können, ohne Patientinnen und Patienten zu gefährden.

fit4translation - ist ein Modul-2b-Projekt der Medizininformatik-Initiative (MII). Es unterstützt akademische Forschende bei der Entwicklung von Medizin-Software, In-Vitro Diagnostika und Prüfprodukten, indem es Lehrmaterialien, Workshops, Beratung und Supportstrukturen bereitstellt, um regulatorische Anforderungen wie MDR, IVDR und Fragen etwa zur Herstellerverantwortung und technischen Dokumentation zu erfüllen. Zudem wird ein Simulations-



**medizinischer
fakultätentag**

Der **Medizinische Fakultätentag (MFT)** ist der Verband der Medizinischen Ausbildungs- und Forschungsstätten Deutschlands. Er vertritt 40 Mitglieder. Diese verantworten Lehre und Forschung in der Human- und Zahnmedizin sowie den Gesundheitswissenschaften. In über 70 verschiedenen Studiengängen mit insgesamt über 120.000 Studierenden. Mit rund 4.000 Professuren und 40.000 Ärzt:innen und Forschenden haben die Medizinischen Fakultäten eine Schlüsselposition in der Biomedizinischen Forschung in Deutschland.

Der **Medizinische Fakultätentag (MFT)** ist der Verband der Medizinischen Ausbildungs- und Forschungsstätten Deutschlands. Er vertritt 40 Mitglieder. Diese verantworten Lehre und Forschung in der Human- und Zahnmedizin sowie den Gesundheitswissenschaften. In über 70 verschiedenen Studiengängen mit insgesamt über 120.000 Studierenden. Mit rund 4.000 Professuren und 40.000 Ärzt:innen und Forschenden haben die Medizinischen Fakultäten eine Schlüsselposition in der Biomedizinischen Forschung in Deutschland.



**VERBAND DER
UNIVERSITÄTSKLINIKA
DEUTSCHLANDS**

Der **Verband der Universitätsklinika Deutschlands e.V. (VUD)** repräsentiert die 37 deutschen Universitätsklinika. Diese sind eine tragende Säule des Gesundheitssystems und stehen für eine Krankenversorgung auf höchstem

Niveau sowie für Spitzenforschung und die Einführung neuer Behandlungsmethoden. Zudem sichern

Revision der MDR / IVDR

Kommentierung und Vorschläge aus der Sicht
der akademischen Forschung und
akademischen Verbundforschung



fit4translation

die Universitätsklinika, gemeinsam mit den medizinischen Fakultäten, die Ausbildung künftiger Generationen von Ärzten und Wissenschaftlern.



Der **Berufsverband Deutscher Anästhesistinnen und Anästhesisten (BDA)** steht mit seinen mehr als 20.000 Mitgliedern für eine starke anästhesiologische ärztliche Gemeinschaft

in den fünf Fachbereichen der Anästhesiologie: Anästhesie, Intensivmedizin, Notfallmedizin, Schmerzmedizin und Palliativmedizin

Der BDA ist die Interessensvertretung für alle beruflichen Belange der Anästhesistinnen und Anästhesisten in sämtlichen Versorgungs- und Fachbereichen der Anästhesiologie auf nationaler und europäischer Ebene. Als starker Partner bietet der BDA seinen Mitgliedern ein umfangreiches Beratungs-, Fortbildungs- und Serviceangebot.



**BERUFSVERBAND
DEUTSCHER PATHOLOGINNEN
UND PATHOLOGEN e.V.**

Der **BDP** ist die berufsständische Vertretung der Fachgebiete Pathologie und Neuropathologie und einer der ältesten Berufsverbände in Deutschland. Er setzt sich seit jeher für gute berufspolitische Rahmenbedingungen für das Fachgebiet ein. Der BDP entwickelt den Beruf

konsequent weiter – und stellt damit die Weichen für eine stets hochqualitative, verlässliche und moderne pathologische Diagnostik. Diese Aufgaben erfüllt er mit seinen über 1.600 Mitgliedern und seinen Mitarbeiterinnen und Mitarbeitern in seiner Geschäftsstelle in Berlin.



Deutsche Gesellschaft
für Anästhesiologie
und Intensivmedizin

Die **"Deutsche Gesellschaft für Anästhesiologie und Intensivmedizin e.V." (DGAI)**, vereinigt als medizinisch-

wissenschaftliche Fachgesellschaft mehr als 15.100 Mitglieder, die deutschlandweit jährlich mehr als 10 Mio. Patientinnen und Patienten aller Altersklassen in den 5 medizinischen Bereichen der Anästhesiologie betreut: Anästhesie, Intensivmedizin, Notfallmedizin, Schmerzmedizin und Palliativmedizin.



Die **DGINA Deutsche Gesellschaft für Notfallmedizin** ist eine medizinische Fachgesellschaft, die sich insbesondere für die Weiterentwicklung der Notfallmedizin und der Notfallpflege einsetzt. Sie sieht die Notfallmedizin als eine klinische Disziplin, die hauptsächlich in den Notaufnahmen betrieben wird, aber auch die präklinische Notfall-, die Katastrophen-

und die Akutmedizin umfasst. Vereinsziele sind vor allem die Verbesserung der notfallmedizinischen Versorgung für Patienten und die Verbesserung der Arbeitsbedingungen in der Notfallmedizin.



Die **Deutsche Gesellschaft für Pathologie (DGP)** ist die wissenschaftliche Fachgesellschaft in der Pathologie und fördert die ärztlichen Belange in dem Bestreben, der Erforschung und Abwehr von Krankheiten zu dienen. Darüber hinaus entwickelt sie die Pathologie in ihrer zentralen Bedeutung für die gesamte Medizin weiter. Die DGP organisiert interdisziplinär ausgerichtete Tagungen und informiert über neueste Erkenntnisse aus Theorie und Praxis der Pathologie. Mit ihren über 1.000 Mitgliedern bietet sie eine Plattform zur Orientierung und zum wissenschaftlichen Austausch in den derzeit 15 Arbeitsgemeinschaften.



Die **Deutsche Gesellschaft für Mund-, Kiefer- und Gesichtschirurgie (DGMKG)** ist der Gesamtverband aller Fachärzte für Mund-, Kiefer- und Gesichtschirurgie in Deutschland. Der Gesamtverband mit über 2.100 Mitgliedern ist ein freiwilliger Zusammenschluss dieser Ärzte zur Wahrung, Förderung und Vertretung der wissenschaftlichen, berufspolitischen, wirtschaftlichen und sonstigen gemeinsamen Belange. Der Zweck

ist die einheitliche und wirkungsvolle Vertretung des Fachgebietes nach innen und außen in Belangen der wissenschaftlichen Darstellung, der berufspolitischen Fragen und der Weiterentwicklung des Fachgebietes in Klinik und Praxis.

Revision der MDR / IVDR

Kommentierung und Vorschläge aus der Sicht
der akademischen Forschung und
akademischen Verbundforschung



fit4translation



Die **Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e. V. (GMDS)** vereint als einzige wissenschaftliche Fachgesellschaft in Deutschland, Österreich und der Schweiz knapp 2.000 Mitglieder aus den Disziplinen der

Medical Data Sciences: Medizinische Informatik, Biometrie, Epidemiologie, Bioinformatik und Systembiologie. Die GMDS fördert interdisziplinär und interprofessionell den Austausch sowie die methodische Weiterentwicklung datenwissenschaftlicher Verfahren in der Medizin. Sie engagiert sich aktiv in gesundheits- und wissenschaftspolitischen Diskursen und gestaltet so maßgeblich die digitale Medizin in Deutschland mit.



Die **Deutsche Gesellschaft für Telemedizin e. V. (DGTelemed)** versteht sich als Forum für Kommunikation, Diskussion und Interessenvertretung in der Telemedizin in Deutschland und Europa. Patientenorientierung und Optimierung der

Zusammenarbeit zwischen Leistungserbringern, Gesundheitsdienstleistungs- und Medizintechnikunternehmen, Verbänden und Vereinigungen gehören zum Grundverständnis ihres Handelns. In ihrem Netzwerk vielfältiger Akteur:innen engagiert sich die DGTelemed für eine optimale Gesundheitsversorgung. Sie nutzt ihr Know-how, um das Bewusstsein für die vielfältigen Möglichkeiten der Telemedizin



Die **Deutsche Interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin e. V. (DIVI)** vereint rund 4.000 Mitglieder aus allen relevanten medizinischen Fachrichtungen und Gesundheitsberufen. Als interdisziplinäre und interprofessionelle

Fachgesellschaft setzt sie sich für eine bestmögliche Versorgung kritisch kranker und verletzter Menschen ein – von der Notfallaufnahme bis zur Intensivstation. Seit vielen Jahren gestaltet die DIVI aktiv die digitale Transformation in der Intensiv- und Notfallmedizin mit. Sie engagiert sich dabei besonders für Interoperabilität medizinischer Systeme und die Digitalisierung der Medizintechnik, um die Versorgungsqualität, Patientensicherheit und Effizienz in der Akutmedizin nachhaltig zu verbessern.



Center for Translational & Clinical Research Aachen (CTC-A)

Das **Center for Translational & Clinical Research Aachen (CTC-A)** an der Uniklinik RWTH Aachen unterstützt mit dem Translationszentrum die akademische Forschung bei der Entwicklung und Herstellung, von Arzneimitteln, innovativen Medizinprodukten und neuen In-vitro-Diagnostika. Das Studienzentrum am CTC-A unterstützt die forschenden Ärztinnen und Ärzte bei der Planung und Durchführung von klinischen Prüfungen. Mit seiner

Infrastruktur begleitet das CTC-A damit den gesamten Prozess von der präklinischen Forschung über die regulatorische Beratung (MDR/IVDR) bis hin zur klinischen Anwendung am Patienten. So schafft das CTC-A eine entscheidende Schnittstelle zwischen Wissenschaft, regulatorischen Anforderungen und klinischer Praxis.



Der **AKTIN e.V. - Aktionsbündnis zur Verbesserung der Kommunikations- und Informationstechnologie in der Intensiv- und Notfallmedizin e.V** ist ein gemeinnütziger Verein, der verschiedene Akteure aus Klinik, Wissenschaft und medizinischer Informatik vereint, um die klinische Notfallversorgung durch verbesserte digitale Routinedaten zu optimieren. Die Vision des Vereins ist eine optimale Notfallversorgung durch Smart Data,

während die Mission darin besteht, Forschung und Versorgung durch effiziente und qualitätsgesicherte Nutzung digitaler Routinedaten aus der Notfall- und Akutversorgung zu verbessern. Dazu werden auch digitale Innovationen in der Notfall- und Akutversorgung erstellt und evaluiert.



Der „**Bonn Hub for Algorithmic Innovation in Medicine**“ (**BoHAIME**) ist ein Fachinkubator für algorithmische Innovationen an der Medizinischen Fakultät der Universität Bonn und dem Universitätsklinikum Bonn. BoHAIME bietet Gründerprojekten und Start-ups am Standort Bonn einen unentgeltlichen Beratungsservice

an, sowie die Möglichkeit ein EN ISO 13485-konformes QMS einzusetzen. Fokus der individuellen Beratung liegt in den Bereichen Qualitäts- und Risikomanagement, Regulatorik und Software Life-Cycle Management. Der BoHAIME-Hub ergänzt die Förderungsstrukturen des Transfer Centers EnaCom für die Start-up Beratung an der Universität Bonn und konnte dank Unterstützung des Ministeriums für Wirtschaft, Industrie, Klimaschutz und Energie des Landes Nordrhein-Westfalen (MWIKE) aufgebaut werden.

Revision der MDR / IVDR

Kommentierung und Vorschläge aus der Sicht
der akademischen Forschung und
akademischen Verbundforschung



fit4translation



Das **QUEST Center for Responsible Research** am Berlin Institute of Health (BIH) ist eine Forschungseinheit der Charité, die darauf abzielt, die Qualität, Ethik und Glaubwürdigkeit der biomedizinischen Forschung zu stärken. Es entwickelt und implementiert u.a. Maßnahmen für mehr Transparenz (z. B. Open Data), Reproduzierbarkeit und Validität von Studien sowie ethisch fundierte Standards bei Forschung mit Menschen, Tieren und sensiblen Daten.