



Marktzugang von kontinuierlich-lernenden KI-Systemen in der Medizin

VDE DGBMT

Marktzugang von kontinuierlich-lernenden KI-Systemen in der Medizin

Autor*innen:

Dr. Thorsten Prinz, VDE e.V.

Dr. Zeynep Schreitmüller, Produktkanzlei Augsburg

Dr. Boris Handorn, Produktkanzlei Augsburg

Herausgeber/Bezugsquelle

VDE Verband der Elektrotechnik Elektronik Informationstechnik e.V.

DGBMT Deutsche Gesellschaft für Biomedizinische Technik im VDE (VDE DGBMT)

Fachausschuss Regulatory Affairs

Merianstraße 28

63069 Offenbach am Main

dgbmt@vde.com

<https://www.vde.com/de/dgbmt>

Bildnachweis:

Titelseite: © catalin/stock.adobe.com

Design: Schaper Kommunikation, Bad Nauheim

Juli 2023/Version 1.1

Inhaltsverzeichnis

Zusammenfassung	4
Abkürzungen und Definitionen	5
Kontinuierlich-lernende KI-Systeme in der Medizin	6
Gesetzliche und normative Anforderungen bei Änderungen an CE-gekennzeichneten Medizinprodukten	6
Orientierungshilfen zur Einstufung von Änderungen	8
Neue regulatorische Ansätze	9
Lösungsansätze für die CE-Konformitätsbewertung von kontinuierlich-lernenden KI-Systemen unter dem geltenden Rechtsrahmen	10
Praktische Umsetzung einer antizipierenden CE-Konformitätsbewertung	10
Beispiel für die Anwendung einer antizipierenden CE-Konformitätsbewertung	11
Empfehlungen	12
Literatur	12
Anhang	14
Struktur eines Predetermined Change Control Plan (PCCP) für die antizipierende CE-Konformitätsbewertung	14

Zusammenfassung

Systeme der künstlichen Intelligenz (KI-Systeme) bieten zum Nutzen von Patient*innen völlig neue Möglichkeiten für die Diagnose und Therapie von Krankheiten. In den vergangenen Jahren erhielt eine wachsende Zahl von KI-Systemen als Medizinprodukt Zugang zu regulierten Märkten wie Europa und den USA. Allerdings handelte es sich hierbei in der Regel um statische KI-Systeme, bei denen der Lernprozess vor der Inbetriebnahme abgeschlossen ist und sich das zugrundeliegende KI-Modell nicht mehr ändert. Hierdurch soll sichergestellt werden, dass sich die Verifizierung und Validierung auf einen bestimmten Entwicklungsstand des KI-Systems beziehen. Durch diese Vorgehensweise werden kontinuierlich-lernende KI-Systeme vom Marktzugang ausgeschlossen und damit die Nutzung eines fortlaufenden Lernprozesses als wesentlicher technischer Vorteil von KI-Systemen verhindert.

Ziel dieser VDE-DGBMT Empfehlung ist es, vor dem Hintergrund des zukünftigen europäischen Artificial Intelligence Act (AIA) und der gegenwärtigen Gesetzgebung zu Medizinprodukten eine Lösung für die bisherige innovationshemmende Vorgehensweise für kontinuierlich-lernende KI-Systeme im europäischen Markt aufzuzeigen. Das Kernstück bildet hierfür eine sogenannte „antizipierende CE-Konformitätsbewertung“, welche die Planung und Genehmigung von beabsichtigten Änderungen bereits vor der Inbetriebnahme vorsieht.

Die VDE-DGMBT Empfehlung richtet sich primär an die nationalen zuständigen Behörden und Benannten Stellen sowie an den Europäischen Gesetzgeber.

Abkürzungen und Definitionen

- **Bias:** Systematic difference in treatment of certain objects, people, or groups in comparison to others (1)
- **Daten-Drift / data drift:** Accuracy of the model's predictions decays over time due to changes in the statistical characteristics of the production data (e.g. image resolution has changed, or one class has become more frequent in data than another) (1)
- **Gesetz über künstliche Intelligenz / artificial intelligence act (AIA):** Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung bestimmter Rechtsakte der Union
- **Kontinuierliches Lernen / continuous learning:** Incremental training of an AI system that takes place on an ongoing basis during the operation phase of the AI system life cycle (1)
- **Künstliche Intelligenz (KI) / artificial intelligence (AI):** Set of methods or automated entities that together build, optimize and apply a model so that the system can, for a given set of predefined tasks, compute predictions, recommendations, or decisions (1)
- **KI-System / AI system:** Engineered system featuring AI (1)
- **Maschinelles Lernen (ML) / machine learning (ML):** Process of optimizing model parameters through computational techniques, such that the model's behaviour reflects the data or experience (1)
- **Machine Learning-Enabled Device Software Function (ML-DSF):** A device software function that implements an ML model trained with ML techniques (2)
- **Medizinprodukte-Verordnung (MP-VO) / Medical Device Regulation (MDR):** Verordnung (EU) 2017/745 des Europäischen Parlaments und des 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
- **Medizinprodukte-Software (MPSW) / medical device software (MDSW):** Medical device software is software that is intended to be used, alone or in combination, for a purpose as specified in the definition of a "medical device" in the medical devices regulation or in vitro diagnostic medical devices regulation (3)
- **Overfitting:** Generation of a ML model that corresponds too closely to the training data, resulting in a model that finds it difficult to generalize to new data (4)
- **Qualitätsmanagementsystem (QMS) / quality management system (QMS):** part of a management system with regard to quality (5)

Kontinuierlich-lernende KI-Systeme in der Medizin

Systeme der künstlichen Intelligenz (KI-Systeme) in der Medizin gewinnen in den letzten Jahren zunehmend an Bedeutung. Im US-amerikanischen und europäischen Markt werden vermehrt neue KI-Systeme als Medizinprodukte zugelassen bzw. in Verkehr gebracht (6,7).

Ein wesentlicher Bestandteil der Entwicklung von KI-Systemen ist deren Training mit geeigneten Trainingsdaten. Dabei kann die Vorgehensweise unterschiedlich sein: In einem Fall werden die Modelle von KI-Systemen während der Entwicklung trainiert und getestet, bevor sie in die Marktphase¹ übergehen und während dieser unverändert bleiben (statische KI-Systeme²). In einem anderen Fall werden Modelle kontinuierlich-lernender KI-Systeme (auch als dynamische KI-Systeme bezeichnet³) in der Marktphase mit neuen Daten mit dem Ziel weiter trainiert, die Leistung des KI-Modells zu verbessern. Durch das kontinuierliche Lernen können außerdem Limitierungen bei der ursprünglichen Verfügbarkeit von Daten überwunden sowie eine breitere Anwendbarkeit des Modells auf weitere Eingangsdaten und eine höhere Generalisierbarkeit erreicht werden.

Es ergibt sich das Problem, dass kontinuierlich-lernende KI-Systeme keinen fixierten technischen Entwicklungsstand aufweisen, da sich dieser mit Blick auf das zugrundeliegende Modell während der Marktphase ändert. Diese Tatsache hat dazu geführt, dass kontinuierlich-lernende KI-Systeme für medizinische Anwendungen derzeit nicht als Medizinprodukte auf dem europäischen Unionsmarkt in Verkehr gebracht und vermarktet werden dürfen.

Gesetzliche und normative Anforderungen bei Änderungen an CE-gekennzeichneten Medizinprodukten

Medizinprodukte unterliegen in Europa den Anforderungen der Verordnung (EU) 2017/745 (im Weiteren bezeichnet als MDR, für engl. *Medical Device Regulation*) (8) bzw. als In-vitro Diagnostika der Verordnung (EU) 2017/746³ (im Weiteren bezeichnet als IVDR, für engl. *In vitro Diagnostic Regulation*). Vor dem Inverkehrbringen müssen Hersteller von Medizinprodukten die Konformität ihrer Produkte mit den Anforderungen der MDR bzw. der IVDR nachweisen. Nach Abschluss des Konformitätsbewertungsverfahrens stellen die Hersteller eine EU-Konformitätserklärung aus und versehen das Produkt mit einer entsprechenden CE-Kennzeichnung.

Nach Anhang I Kapitel II Abschnitt 17.1 MDR ist Software so auszulegen, dass Wiederholbarkeit, Zuverlässigkeit und Leistung entsprechend ihrer bestimmungsgemäßen Verwendung gewährleistet werden. Die CE-Konformitätsbewertung und das nachfolgende Inverkehrbringen sind daher in der Regel an einen definierten technischen Entwicklungsstand gekoppelt. Im weiteren Lebenszyklus unterliegen Medizinprodukte jedoch im Regelfall technischen Änderungen, die je nach Umfang und Kritikalität auch eine erneute CE-Konformitätsbewertung erforderlich machen können.

¹ Der Begriff „Marktphase“ bezieht sich hier auf das Inverkehrbringen und den Betrieb von KI-Systemen. In der technischen Community wird diese auch als „Produktionsphase“ bezeichnet.

² Vgl. German Notified Bodies Alliance, Hrsg., Questionnaire „Artificial Intelligence (AI) in medical devices“.

³ Da KI-Systeme in der Medizin aufgrund ihrer Zweckbestimmung in der Regel den Anforderungen der MDR unterliegen, werden in diesem Dokument Anwendungsfälle als In-vitro Diagnostika nicht betrachtet.

Angesichts der praktischen Relevanz hat der Unionsgesetzgeber daher den Umgang mit Änderungen an Medizinprodukten mehrfach in der MDR normiert. So fordert Art. 10 Abs. 9 UAbs. 1 S. 2 MDR, dass ein Hersteller „Änderungen an der Auslegung des Produkts oder an seinen Merkmalen [...] zeitgerecht angemessen berücksichtigt“. Im Qualitätsmanagementsystem (QMS) des Herstellers wird die Kontrolle von Design- und Entwicklungsänderungen in einem entsprechenden Prozess umgesetzt (Art. 10 Abs. 9 UAbs. 3 lit. a MDR).

Für den Fall, dass Hersteller die Konformitätsbewertung auf der Grundlage eines QMS und einer Bewertung der technischen Dokumentation nach Anhang IX MDR durchführen, müssen Änderungen an einem genehmigten Produkt von der Benannten Stelle, die die EU-Bescheinigung über die Bewertung der technischen Dokumentation ausgestellt hat, genehmigt werden, wenn diese Änderungen die Sicherheit und Leistungsfähigkeit des Produkts oder die für das Produkt vorgeschriebene Anwendungsbedingungen beeinträchtigen könnten (Anhang IX Kapitel II Abschnitt 4.10. S. 1 MDR). Eine vergleichbare Bestimmung findet sich in Anhang X MDR für die Konformitätsbewertung auf der Grundlage einer Baumusterprüfung. Gemäß Anhang X Abschnitt 5.2. MDR sind Änderungen am genehmigten Produkt, einschließlich Beschränkungen seiner Zweckbestimmung oder seiner Verwendungsbedingungen, von der Benannten Stelle, die die Prüfbescheinigung ausgestellt hat, zu genehmigen, wenn diese Änderung die Konformität des Produkts mit den grundlegenden Sicherheits- und Leistungsanforderungen oder mit den vorgesehenen Anwendungsbedingungen des Produkts beeinträchtigen können. Benannte Stellen müssen Herstellern in diesen Fällen ein Verfahren zur Bewertung von Änderungen zur Verfügung stellen (Abschnitt 4.9. Anhang VII MDR). Basierend auf dem Verfahren hat die Benannte Stelle zu entscheiden, ob die jeweiligen Änderungen eine neue CE-Konformitätsbewertung erfordern.

Ferner sieht die MDR im Falle von Änderungen der Zweckbestimmung und der Verwendungsbedingungen des genehmigten Produkts vor, dass eine neue Konformitätsbewertung durchgeführt wird (Anhang X Abschnitt 5.3. MDR).

Diese Vorgehensweise hat auch Eingang in technische Normen gefunden. Gemäß Abschnitt 7.3.9 zur Lenkung der Entwicklungsänderungen in der EN ISO 13485 müssen Änderungen genehmigt und hinsichtlich ihres Einflusses auf die Zweckbestimmung sowie auf die Sicherheit, Leistungsfähigkeit und die anwendbaren regulatorischen Anforderungen bewertet werden. Weiterhin muss festgelegt werden, mithilfe welcher Verifizierungs- und Validierungsaktivitäten überprüft wird, dass die Änderungen wie geplant umgesetzt wurden. Hier besteht außerdem eine Rückkopplung zum Risikomanagement. In Bezug auf Medizinprodukte-Software werden Änderungen im Software-Wartungs-Prozess geplant, umgesetzt und kontrolliert.

An den genannten Bestimmungen wird ersichtlich, dass nicht jede Änderung zu einer Neubewertung des Medizinprodukts führt. Dies wird insbesondere auch mit Blick auf Anhang IX Kapitel I Abschnitt 2.4. S. 4 MDR ersichtlich, wonach „nur wesentliche Änderungen“ am Qualitätsmanagement oder der hiervon erfassten Produktpalette durch die Benannte Stelle berücksichtigt werden.

Orientierungshilfen zur Einstufung von Änderungen

Die *Notified Body Operations Group* (NBOG) hat 2014 unter dem alten europäischen Rechtsrahmen in einem Leitfaden Empfehlungen veröffentlicht, wie verschiedene Änderungen in Hinsicht auf die CE-Konformitätsbewertung einzuschätzen sind (9). Der NBOG-Leitfaden nennt die folgenden Beispiele für **wesentliche Änderungen** von Medizinproduktesoftware:

- Änderung mit Auswirkungen auf die Gerätesteuerung,
- Änderung des Algorithmus mit Auswirkungen auf Diagnose oder Therapie,
- Änderungen mit Auswirkung auf die Art und Weise der Dateninterpretation durch den Anwender,
- Ersatz von Benutzereingaben durch *closed-loop*-Eingaben,
- Einführung einer neuen wesentlichen Softwarefunktion,
- Einführung oder Entfernung einer Alarmfunktion sowie
- signifikante Änderung des Betriebssystems

Dem gegenüber stehen **nicht-wesentliche Änderungen** wie bspw. die Einführung neuer, für die Zweckbestimmung nicht wesentlicher Softwarefunktion(en), die geringfügige Änderung des Aussehens der Benutzeroberfläche oder die Deaktivierung einer Softwarefunktion. Eine ähnliche Herangehensweise zur Abgrenzung hinsichtlich der Kritikalität von Änderungen findet sich in der Leitlinie MDCG 2020-3 (10).

In Analogie zum NBOG-Dokument beschreiben Zinchenko und Co-Autoren in einem Artikel, wann Änderungen eines KI-Modells als signifikant und nicht-signifikant anzusehen sind (11). Als Beispiele für **signifikante Änderungen** nennen die Autoren Änderungen in Bezug auf:

- die Effizienz und die Sicherheit (z. B. Verbesserung der Verarbeitungszeit / Erkennungsrate oder Funktionserweiterungen wie *saliency maps*),
- die Eingabedaten (z. B. Kompatibilität mit neuen Herstellerdaten oder zusätzliche Eingabedaten, beides ohne Änderung der Zweckbestimmung) sowie
- die Funktionalitäten (z. B. neue Patientenzielgruppen und Anwendungsbedingungen, beides mit Änderung der Zweckbestimmung).

Nicht-signifikante Änderungen ohne Einfluss auf die Zweckbestimmung betreffen:

- Fehlerbeseitigungen (z. B. Änderung der Anzeige der Ausgabedaten / des Berichtsformats),
- Fehlerberichte und Protokolle (z. B. Änderungen in der Darstellung) oder
- die Benutzeroberfläche (z. B. Änderungen der Fenstergrößen oder Farben).

Die in den beiden Dokumenten beschriebene Sicht auf die Signifikanz von Änderungen impliziert, dass kontinuierlich-lernende KI-Systeme nicht gesetzeskonform sein können, weil fortlaufende Änderungen in den meisten Fällen neue CE-Konformitätsbewertungen erfordern würden.

Die Interessengemeinschaft der Benannten Stellen für Medizinprodukte in Deutschland (IG-NB) äußert sich im Abschnitt A des „Questionnaire Artificial Intelligence (AI) in medical devices“ wie folgt zur Zertifizierbarkeit von KI-Systemen: „Static AI (AI that has learned and operates in a learned state) is in principle certifiable. Dynamic AI (AI that continues to learn in the field) is not certifiable in principle, as the

system must be verified and validated (among other things, the functionality must be validated against the intended use)" (12).

Nach Ansicht der IG-NB muss also ein definierter Entwicklungsstand zum Zeitpunkt der Verifizierungs- und Validierungsaktivitäten gegeben sein⁴. Demzufolge sind kontinuierlich-lernende KI-Systeme in der Medizin, die durch das kontinuierliche Lernen keinen definierten technischen Entwicklungsstand aufweisen, nach Ansicht der deutschen Benannten Stellen grundsätzlich nicht zertifizierbar.

Neue regulatorische Ansätze

Die FDA hat 2019 in einem Diskussionspapier einen neuen regulatorischen Rahmen für den Marktzugang von KI-Systemen in der Medizin vorgeschlagen (13). So sollen die Sicherheit und Leistungsfähigkeit dieser KI-Systeme durch ein umfassendes QMS auf der Basis von *Good Machine Learning Practices* (GMLP) sichergestellt werden. 2021 veröffentlichten FDA, Health Canada und die englische Medicines and Healthcare products Regulatory Agency (MHRA) einen entsprechenden Leitfaden (14). Unter bestimmten Voraussetzungen sieht das FDA-Diskussionspapier auch für kontinuierlich-lernende KI-Systeme den Marktzugang vor. Kernstück dieses neuen regulatorischen Ansatzes ist der sog. **Predetermined Change Control Plan (PCCP)**, in dem der Hersteller vor dem Marktzugang geplante zukünftige Änderungen am KI-Modell und deren Evaluierung offenlegen soll, wie die FDA kürzlich in einem Leitfaden-Entwurf ausführlich dargestellt hat (2).

Der PCCP besteht aus einer detaillierten **Beschreibung der Änderungen**, einem **Änderungsprotokoll** und einer **Folgenabschätzung**. In der detaillierten Beschreibung sollen die Modifikationen der Geräteeigenschaften und -leistungen beschrieben werden, die sich aufgrund der Änderungen ergeben. Das Änderungsprotokoll umfasst die Beschreibung der Umsetzung der geplanten Änderungen sowie der damit verbundenen Verifizierungs- und Validierungsaktivitäten. Abschließend soll die Folgenabschätzung eine Bewertung des Nutzens und der Risiken der geplanten Änderungen vornehmen. Der PCCP soll von der FDA vor Marktzulassung geprüft und genehmigt werden. Solange sich die Änderungen des KI-Systems in der Marktphase im Rahmen der angezeigten und genehmigten Änderungen bewegen, muss kein neues Zulassungsverfahren durchlaufen werden.

Auch auf europäischer Ebene sind in den letzten Jahren KI-spezifische Regulierungsaktivitäten zu verzeichnen. Die EU-Kommission hat 2021 den Entwurf eines europäischen *Artificial Intelligence Act* (AIA-E) veröffentlicht (15). KI-Systeme als Medizinprodukte werden darin als sog. Hochrisiko-KI-Systeme eingestuft und unterliegen besonderen Marktzutrittsanforderungen. Grundsätzlich sieht der AIA-E als Entwurf einer EU-Verordnung vor, dass Hochrisiko-KI-Systeme einem neuen CE-Konformitätsbewertungsverfahren unterzogen werden, wenn diese „wesentlich geändert werden, unabhängig davon, ob das geänderte System noch weiter in Verkehr gebracht oder vom derzeitigen Nutzer weitergenutzt werden soll“ (Art. 43 Abs. 4 UAbs. 1 AIA-E). Gemäß Erwägungsgrund Nr. 66 AIA-E handelt es sich um eine **wesentliche Änderung**, wenn mit ihr die Einhaltung des AIA beeinträchtigt werden könnte oder wenn sie die Zweckbestimmung des Systems ändert.

Ähnlich dem Ansatz im FDA-Diskussionspapier werden hingegen „Änderungen des Hochrisiko-KI-Systems und seiner Leistung, die vom Anbieter zum Zeitpunkt der ursprünglichen CE-Konformitätsbewertung vorab festgelegt wurden und in den Informationen der technischen Dokumentation gemäß Anhang IV Nummer 2 Buchstabe f enthalten sind, nicht als wesentliche Änderung“ angesehen (Art. 43 Abs. 4 UAbs. 2 AIA-E). Dieser Gesetzesentwurf des europäischen Gesetzgebers ist deutlich später als die MDR erschienen und spiegelt deshalb auch eine regulatorische Entwicklung vor dem Hintergrund neuer technologischer Entwicklungen wider.

Der europäische Gesetzgeber würde mit dem AIA eine völlig neue Art der CE-Konformitätsbewertung für kontinuierlich-lernende KI-Systeme vorsehen, die bislang so weder von den deutschen Benannten Stellen noch in der MDR explizit vorgesehen ist.

⁴ In diesem Dokument werden die Begriffe „Verifizierung“ und „Validierung“ gemäß den Definitionen in der ISO 9000 angewendet (5).

Lösungsansätze für die CE-Konformitätsbewertung von kontinuierlich-lernenden KI-Systemen unter dem geltenden Rechtsrahmen

Überträgt man die beiden zuvor genannten regulatorischen Ansätze, könnte in Europa eine sogenannte „antizipierende CE-Konformitätsbewertung“ für kontinuierlich-lernende KI-Systeme in der Medizin als Teil des geltenden Rechtsrahmens eingeführt werden (16). Die antizipierende CE-Konformitätsbewertung wäre dadurch gekennzeichnet, dass diese vorgreifend, unter Einschluss der beabsichtigten Änderungen im Laufe der Inbetriebnahme, erfolgt. Für spätere Änderungen, die sich im Rahmen des Antizipierten bewegen, könnte von einer weiteren Konformitätsbewertung abgesehen werden. Änderungen, die nicht vorhersehbar und denknotwendig nicht antizipiert werden können, wären dann einem erneuten Konformitätsbewertungsverfahren zu unterziehen und nachträglich zu zertifizieren (16).

Auch wenn dieser Ansatz im europäischen Medizinproduktrecht eine Neuerung darstellen würde, ist dieser mit dem geltenden Rechtsrahmen der MDR vereinbar. Dafür spricht bereits der Umstand, dass die oben referenzierten Vorschriften (vgl. Abschnitt "Gesetzliche und normative Anforderungen bei Änderungen an CE-gekennzeichneten Medizinprodukten") offen gefasst sind. Eine vorgreifende Konformitätsbewertung wird nicht explizit ausgeschlossen. Dieser steht insbesondere nicht entgegen, dass Anhang I Kapitel II Abschnitt 17.1. MDR fordert, dass die Software so auszulegen ist, dass Wiederholbarkeit gewährleistet ist. Das Qualitätsmerkmal der Wiederholbarkeit bezieht sich nach dem Sinn und Zweck nicht darauf, das System möglichst statisch auszugestalten. Andernfalls würden auch notwendige Software-Updates oder Patches dazu führen, dass ein initial bewertetes System nicht mehr wiederholbar im Sinne der MDR ist. Vielmehr bezieht sich die regulatorische Anforderung im Rahmen der Konformitätsbewertung darauf, dass Wiederholbarkeit aus Gründen der Nachvollziehbarkeit des Outputs in den Phasen zwischen den vorgenommenen Software-Änderungen gewährleistet ist.

Des Weiteren lassen sich die Abschnitte 5.2 und 5.3 des Anhangs X MDR dahingehend auslegen, dass antizipierte Änderungen, die bereits in einem initialen Konformitätsbewertungsverfahren geprüft wurden, den Schutzzweck der MDR ebenso erfüllen wie ein nachträglich durchgeführtes Bewertungsverfahren. Schließlich ist Sinn und Zweck einer erneuten Konformitätsbewertung eine Neubewertung der mit der relevanten Änderung verbundenen Risiken für die Patientensicherheit. Diese würde hinfällig werden, würde man die Bewertung bereits beim ersten Durchlauf erfassen und entsprechend bewerten. Insgesamt ließe sich der Ansatz der FDA und der EU-Kommission daher regulatorisch auch unter der geltenden MDR realisieren.

Praktische Umsetzung einer antizipierenden CE-Konformitätsbewertung

Zum Zeitpunkt der Konformitätsbewertung könnten durch die Benannte Stelle definierte geplante Änderungen auf Basis eines PCCP als Teil der technischen Dokumentation bereits genehmigt werden. Dadurch muss der Hersteller diese nicht mehr zusätzlich im Prozess „Kontrolle von Design- und Entwicklungsänderungen“ adressieren und nach der Inbetriebnahme der Benannten Stelle melden bzw. sich genehmigen lassen.

Im Rahmen einer antizipierenden CE-Konformitätsbewertung kommt der Überwachung nach dem Inverkehrbringen eine besondere Bedeutung in der Sicherheitsarchitektur kontinuierlich lernender KI-Systeme zu. So heißt es beispielsweise in Art. 61 Abs. 1 AIA-E, dass der Anbieter (Hersteller i. S. d. MDR) ein System zur Beobachtung nach dem Inverkehrbringen einrichten und dokumentieren muss, dass im Verhältnis zur Art der KI-Technik und zu den Risiken des Hochrisiko-KI-Systems steht. Im Erwägungsgrund Nr. 54 des Kommissionsentwurfs wird ergänzt, dass es sich hierbei um

ein "robustes" System zur Beobachtung nach dem Inverkehrbringen handeln soll. Weiterhin sieht die GMLP-Richtlinie vor, dass die Leistung von KI-Modellen in der Marktphase überwacht und die Risiken durch Re-Training gemanagt werden (14). In Bezug auf kontinuierlich-lernende KI-Systeme sollen geeignete Kontrollen zur Handhabung von Risiken in Bezug auf *Overfitting*, *Bias* oder Modellverschlechterung (z. B. durch *Data Drift*) vom Hersteller implementiert werden (14). Im Rahmen der *Statistical Process Control* (SPC) können verschiedene Formen von statistischen Kontrollkarten beispielsweise dazu dienen, die Leistung der KI-Systeme in der Marktphase zu überwachen (17). Diesbezüglich können Hersteller auch auf verschiedene Software-Pakete zurückgreifen, z. B. *strucchange*®, *Scikit-multiflow* (Python) oder *Massive Online Analysis* (MOA) (18–20).

Beispiel für die Anwendung einer antizipierenden CE-Konformitätsbewertung

Zur Verdeutlichung erläutern wir die antizipierende CE-Konformitätsbewertung anhand eines konkreten Produktbeispiels.

Als Beispiel dient ein KI-System, welches Angio-CT-Bilder des Gehirns in der Notaufnahme der Klinik automatisiert analysiert. Sobald das KI-System einen potenziellen Verschluss eines großen Gefäßes identifiziert hat, empfiehlt es dem zuständigen Radiologen die Überprüfung dieser Bilder und unterstützt so die Triage parallel zum Standard-Behandlungsablauf in der Klinik.

Als Teil einer antizipierenden CE-Konformitätsbewertung beschreibt der Hersteller im PCCP, dass er in den ersten 6 Monaten nach dem Inverkehrbringen beabsichtigt, Bilder von einem CT-Scanner eines zweiten Herstellers für das kontinuierliche Lernen des KI-Modells zu verwenden. Der Hersteller legt Akzeptanzkriterien hinsichtlich Bildcharakteristika fest und beschreibt, wie die Einhaltung der Akzeptanzkriterien der CT-Bilder überwacht werden soll.

Der Hersteller stellt dar, dass diese Änderung des KI-Systems keine Änderung der Zweckbestimmung erforderlich macht, sondern einer besseren Generalisierbarkeit des KI-Modells dient. Im PCCP gibt der Hersteller an, dass die Bilder des CT-Scanners vom zweiten Hersteller durch Radiologen innerhalb eines festgelegten Zeitraums auf potenzielle Gefäßverschlüsse überprüft werden. Im Rahmen der Überwachung nach dem Inverkehrbringen legt er Akzeptanzkriterien für ausgewählte Gütekriterien in Bezug auf das KI-Modell fest, welche ein fortwährendes Leistungsmonitoring ermöglichen. Der Hersteller beschreibt auch die eingesetzten Software-Tools und implementiert eine automatische Alarm-Funktion sowie eine entsprechende Handlungs-Anweisung, sobald die Gütekriterien unter einen kritischen Wert fallen.

Die Benannte Stelle überprüft im Rahmen der initialen CE-Konformitätsbewertung den PCCP und stellt fest, dass die geplanten Änderungen am KI-System als wesentlich anzusehen sind und deshalb grundsätzlich eine erneute CE-Konformitätsbewertung erfordern. Die Maßnahmen zur Risikokontrolle erscheinen der Benannten Stelle jedoch ausreichend. Die Benannte Stelle genehmigt die im PCCP dargestellten Änderungen.

Würde der Hersteller Angio-CT-Bilder von weiteren Patientengruppen (z. B. in Hinsicht auf das Geschlecht oder die Altersgruppe) für das kontinuierliche Lernen des KI-Modells verwenden und diese wären im PCCP nicht beschrieben und genehmigt worden, wäre grundsätzlich eine erneute CE-Konformitätsbewertung erforderlich.

Empfehlungen

Im Sinne einer antizipierenden CE-Konformitätsbewertung sollten Benannte Stellen einen vom Hersteller eingereichten PCCP für kontinuierlich-lernende KI-Systeme mit geplanten Änderungen am Modell bewerten. Sofern sich die Änderungen innerhalb des genehmigten PCCP bewegen, sollte keine erneute CE-Konformitätsbewertung mit Blick auf Änderungen am Modell durch die Benannte Stelle erforderlich sein. Dies gilt unter der Prämisse, dass ausreichende Maßnahmen zur Minderung etwaiger neuer Risiken und eine effektive Überwachung nach dem Inverkehrbringen implementiert wurden.

Die *Medical Device Coordination Group* der EU-Kommission (MDCG) sollte eine Leitlinie zum Inverkehrbringen von KI-Systemen in der Medizin veröffentlichen und den Benannten Stellen explizite Empfehlungen zum Umgang mit kontinuierlich-lernenden KI-Systemen geben. Dazu gehören insbesondere Vorgaben zu Form und Funktion eines PCCP, die der Hersteller zu beachten hat. Der Anhang dieses Dokumentes enthält einen Gliederungsentwurf für den PCCP, wie dieser für den europäischen Markt eingeführt werden könnte. Weiterhin sollte diese MDCG-Leitlinie nachvollziehbare Beurteilungskriterien für die Benannte Stelle in Bezug auf die Prüfung und Genehmigung eines PCCP enthalten. Die MDCG sollte darauf achten, dass die Erstellung der MDCG-Leitlinie in enger Abstimmung mit den Anforderungen des zukünftigen AIA erfolgt.

Literatur

1. ISO/IEC, Herausgeber. ISO/IEC 22989 Information technology - Artificial intelligence - Artificial intelligence concepts and terminology [Internet]. 2022. Verfügbar unter: <https://www.iso.org/obp/ui#iso:std:iso-iec:22989:ed-1:v1:en>
2. FDA, Herausgeber. Draft Guidance: Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence/Machine Learning (AI/ML)-Enabled Device Software Functions [Internet]. 2023. Verfügbar unter: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial>
3. Medical Device Coordination Group, Herausgeber. MDCG 2019-11 Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 – MDR and Regulation (EU) 2017/746 – IVDR [Internet]. 2019. Verfügbar unter: https://health.ec.europa.eu/medical-devices-sector/new-regulations/guidance-mdcg-endorsed-documents-and-other-guidance_en
4. ISO/IEC, Herausgeber. ISO/IEC TR 29119-11 Software and systems engineering - Software testing - Part 11: Guidelines on the testing of AI-based systems [Internet]. 2020. Verfügbar unter: <https://www.iso.org/obp/ui#iso:std:iso-iec:tr:29119:-11:ed-1:v1:en:term:3.1.12>
5. DIN e.V., Herausgeber. DIN EN ISO 9000 Qualitätsmanagementsysteme - Grundlagen und Begriffe [Internet]. 2015. Verfügbar unter: <https://www.iso.org/obp/ui#iso:std:iso:9000:ed-4:v1:en>
6. FDA. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices [Internet]. 2022 [zitiert 1. Juni 2023]. Verfügbar unter: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>
7. Radboud University Medical Center. AI for Radiology [Internet]. [zitiert 1. Juni 2023]. Verfügbar unter: <http://www.AIforRadiology.com/>

8. Europäisches Parlament und Rat der Europäischen Union. Verordnung (EU) 2017/745 vom 5. April 2017 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 [Internet]. 2017/745 Mai 25, 2017. Verfügbar unter: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02017R0745-20230320>
9. NBOG, Herausgeber. Guidance for manufacturers and Notified Bodies on reporting of Design Changes and Changes of the Quality System [Internet]. 2014. Verfügbar unter: http://www.doks.nbog.eu/Doks/NBOG_BPG_2014_3.pdf
10. Medical Device Coordination Group, Herausgeber. MDCG 2020-3 Guidance on significant changes regarding the transitional provision under Article 120 of the MDR with regard to devices covered by certificates according to MDD or AIMDD [Internet]. 2020. Verfügbar unter: https://health.ec.europa.eu/medical-devices-sector/new-regulations/guidance-mdcg-endorsed-documents-and-other-guidance_en
11. Zinchenko V, Chetverikov S, Akhmad E, Arzamasov K, Vladzmyrskyy A, Andreychenko A, u. a. Changes in software as a medical device based on artificial intelligence technologies. *Int J Comput Assist Radiol Surg.* 1. Oktober 2022;17(10):1969–77.
12. German Notified Bodies Alliance, Herausgeber. Questionnaire „Artificial Intelligence (AI) in medical devices“ (Ver. 4) [Internet]. 2022. Verfügbar unter: <http://www.ig-nb.de/>
13. FDA, Herausgeber. 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 [Internet]. 2019. Verfügbar unter: <https://www.fda.gov/MedicalDevices/DigitalHealth/SoftwareasaMedicalDevice/ucm634612.htm>
14. FDA, Herausgeber. Good Machine Learning Practice for Medical Device Development: Guiding Principles [Internet]. 2021. Verfügbar unter: <https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles>
15. European Commission. Proposal for a Regulation on a European approach for Artificial Intelligence | Shaping Europe's digital future [Internet]. 2021/0106 (COD) 2021. Verfügbar unter: <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-approach-artificial-intelligence>
16. Schreitmüller Z. Regulierung intelligenter Medizinprodukte: eine Analyse unter besonderer Berücksichtigung der MPVO und DSGVO. 1. Auflage. Baden-Baden: Nomos; 2023. 346 S. (Schriften zum Bio-, Gesundheits- und Medizinrecht).
17. Feng J, Phillips RV, Malenica I, Bishara A, Hubbard AE, Celi LA, u. a. Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in health-care. *Npj Digit Med.* 31. Mai 2022;5(1):1–9.
18. Zeileis A, Leisch F, Hornik K, Kleiber C. strucchange: An R Package for Testing for Structural Change in Linear Regression Models. *J Stat Softw.* 10. Januar 2002;7:1–38.
19. Montiel J, Read J, Bifet A, Abdessalem T. Scikit-Multiflow: A Multi-output Streaming Framework [Internet]. 2018 [zitiert 20. Februar 2023]. Verfügbar unter: <http://arxiv.org/abs/1807.04662>
20. Bifet A, Holmes G, Pfahringer B, Kranen P, Kremer H, Jansen T, u. a. MOA: Massive Online Analysis, a Framework for Stream Classification and Clustering. *JMLR Workshop Conf Proc.* 2010;11.

Anhang

Struktur eines Predetermined Change Control Plan (PCCP) für die antizipierende CE-Konformitätsbewertung⁵

A. Description of Modifications

Planned modifications	<p><i>Description of individual proposed device modifications, including:</i></p> <ul style="list-style-type: none">• <i>Types of modifications, e.g., improvement of quality characteristics (e.g., higher sensitivity and specificity), modifications related to device inputs, or modifications related to the device's use and performance (e.g., for use within a specific subpopulation)</i>• <i>Implementation methods, e.g., ML methods, including architecture and parameters, data pre-processing, or both</i>
Impact assessment	<p><i>Discussion of benefits and risks of implemented modifications, including:</i></p> <ul style="list-style-type: none">• <i>Comparison of modified and original device</i>• <i>Appropriateness of the chosen approaches to ensure the safety and effectiveness of the device</i>• <i>Potential interactions between individual modifications given the collective impact of all planned modifications</i>

B. Modification Protocol

B.1 Data management

Person(s) in charge	<p><i>Who is involved in the data management? Name, function (e.g., Data Scientist)</i></p>
Procedure	<p><i>Description of procedure of management of new data, including:</i></p> <ul style="list-style-type: none">• <i>Outline how those new data will be collected, annotated, curated, stored, retained, controlled, and used by the manufacturer for each modification, incl.</i><ul style="list-style-type: none">- <i>Quality assurance (QA) plan for determining which new data are appropriate for inclusion as part of an expanded training data set</i>- <i>data augmentation strategy that allows for additional training and independent test data to be added</i>• <i>Clarification of the relationship between the modification protocol data and the data used to train and test the initial and subsequent versions of the ML-DSF, incl.</i><ul style="list-style-type: none">- <i>Approach to the reference standard determination</i>• <i>Description of control methods preventing data or performance information leaking into the development process during modification development or assessment</i><ul style="list-style-type: none">- <i>strategy to monitor and document test dataset independence as well as control access to both the training and test datasets as additional data are being included and any revised algorithm is being retrained and tested</i>

B.2 Re-training

Person(s) in charge	<p><i>Who is involved in the re-training? Name, function (e.g., ML engineer)</i></p>
Procedure	<p><i>Description of re-training procedure, including:</i></p> <ul style="list-style-type: none">• <i>Outline of re-training strategy that describes the objective of the retraining (e.g., modifications of processing steps and ML architecture)</i>• <i>Training intervals of new algorithm</i>• <i>Modified algorithm components as a result of the learning process</i>• <i>Definition of criteria that must be met during the re-training process to trigger a more comprehensive performance evaluation using the test dataset</i>

⁵ Basierend auf dem FDA Guidance-Entwurf „Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence/Machine Learning (AI/ML)-Enabled Device Software Functions“ (2).

B.3 Performance evaluation

Person(s) in charge	<i>Who is involved in the performance evaluation? Name, function (e.g., ML engineer)</i>
Procedure	<i>Description of procedure for performance evaluation, including:</i> <ul style="list-style-type: none">• <i>Appropriate metrics as performance targets</i>• <i>Test procedures and protocols, which may be applicable for that device and type of change</i>• <i>Intervals for technical evaluation</i>• <i>Definition of- appropriate measures to minimize information leakage about the test data set if part of it is re-used in multiple evaluations</i>• <i>Planned real-world monitoring of performance across data acquisition systems</i>

B.4 Update

Person(s) in charge	<i>Who is involved in the update procedure? Name, function (e.g., ML engineer)</i>
Procedure	<i>Description of update procedure, including:</i> <ul style="list-style-type: none">• <i>Update plan including expected frequency of updates and whether the updates will be global (all devices use the same version of the algorithm) or local (multiple versions of the algorithm targeted for specific sub-populations are distributed)</i>• <i>Version tracking and control</i>• <i>Obsolescence planning</i>• <i>Requirements for host software/hardware requirements</i>• <i>Planned 'beta' release of the updated medical device algorithm concurrent with the previous version</i>

B.5 Communication with users

Person(s) in charge	<i>Who is involved in the communication with users? Name, function (e.g., RA manager)</i>
Procedure	<i>Description of communication procedure with users, including:</i> <ul style="list-style-type: none">• <i>Notification of users of updates and any information that will be conveyed to users about the update (e.g., changes labelling or instruction for use)</i>

Über die Deutsche Gesellschaft für Biomedizinische Technik im VDE (VDE DGBMT)

Die Deutsche Gesellschaft für Biomedizinische Technik im VDE (VDE DGBMT) ist mit über 2.500 Mitgliedern die wissenschaftlich-technische Fachgesellschaft der Medizintechnik in Deutschland. Sie wurde 1961 in Frankfurt am Main gegründet.

Die Vernetzung von Expertinnen und Experten aus dem Bereich der Biomedizinischen Technik ist der Schwerpunkt der Arbeit der VDE DGBMT. Darüber hinaus bietet sie Tagungen und Workshops für Fachpublikum und ist Trägerin von zwei internationalen wissenschaftlichen Zeitschriften: Biomedical Engineering und Current Directions in Biomedical Engineering des Verlags Walter de Gruyter. Positionspapiere, Stellungnahmen und Expertenbeiträge beleuchten unabhängig und neutral aktuelle Themen. Außerdem verleiht die DGBMT Förderpreise für wissenschaftlichen Nachwuchs, für wissenschaftliche Exzellenz und Innovationen und für Patientensicherheit in der Biomedizintechnik. Nicht zuletzt vertritt sie die deutsche Biomedizinische Technik in internationalen Gremien.

Mehr Informationen unter www.vde.com/dgbmt

Über den VDE

Der VDE, eine der größten Technologie-Organisationen Europas, steht seit mehr als 125 Jahren für Innovation und technologischen Fortschritt. Als einzige Organisation weltweit vereint der VDE dabei Wissenschaft, Standardisierung, Prüfung, Zertifizierung und Anwendungsberatung unter einem Dach. Das VDE Zeichen gilt seit mehr als 100 Jahren als Synonym für höchste Sicherheitsstandards und Verbraucherschutz.

Wir setzen uns ein für die Forschungs- und Nachwuchsförderung und für das lebenslange Lernen mit Weiterbildungsangeboten „on the job“. Im VDE Netzwerk engagieren sich über 2.000 Mitarbeiter*innen an über 60 Standorten weltweit, mehr als 100.000 ehrenamtliche Expert*innen und rund 1.500 Unternehmen gestalten im Netzwerk VDE eine lebenswerte Zukunft: vernetzt, digital, elektrisch.

Wir gestalten die e-diale Zukunft.

Sitz des VDE (VDE Verband der Elektrotechnik Elektronik und Informationstechnik e.V.) ist Frankfurt am Main.

Mehr Informationen unter www.vde.com

VDE Verband der Elektrotechnik
Elektronik Informationstechnik e.V.

Merianstraße 28
63069 Offenbach am Main
Tel. +49 69 6308-0
service@vde.com

VDE