

Whitepaper Reha

KI in der Rehabilitation: klinische Qualität, Zusammenarbeit und wirtschaftlicher Nutzen



Dr. med. Sven Jungmann

September 2026

EINFÜHRUNG

Was sich mit KI in der Rehabilitation verändert

Vor einer Reha-Aufnahme müssen Entlassungsbriefe, Medikationspläne und weitere Befunde aus unterschiedlichen Einrichtungen zusammengetragen werden. Während des Aufenthalts kommen eigene Untersuchungen und die Dokumentation mehrerer Berufsgruppen hinzu. KI kann diese Unterlagen aufbereiten und daraus beispielsweise einen Aufnahmebericht oder den Entwurf eines Entlassungsberichts erstellen. Für eine Klinikleitung stellt sich damit eine weitergehende Frage: Was verbessert sich im Betrieb, wenn ein Teil dieser Arbeit automatisiert wird?

Steht dem Aufnahmeteam eine geprüfte Übersicht der Vorbefunde früher zur Verfügung, kann es offene Fragen zur Aufnahme früher klären. Erstellen Stationsärzt:innen ihre Entlassungsentwürfe schneller, können bei den freigebenden Oberärzt:innen mehr Texte gleichzeitig eintreffen. Auch ein sprachlich überzeugender Entwurf kann Angaben aus der Patientenakte auslassen. Ob die Klinik insgesamt entlastet wird, hängt deshalb auch von der anschließenden Prüfung und den übrigen Arbeitsschritten ab.

Dieses Whitepaper richtet sich an Geschäftsführungen, Klinikleitungen und Chefärzt:innen von Rehakliniken sowie Häusern mit Akut- und Rehabilitationsversorgung. Es untersucht, was veröffentlichte Forschung zu Aufnahmeabläufen, KI-gestützter Dokumentation und Zusammenarbeit beitragen kann. Jede Studie wird dort eingeordnet, wo sie eine konkrete Frage beantwortet. Beispiele zeigen, welche Folgen sich daraus für Auswahl und Einführung einer Anwendung ergeben können.

Die Kapitel folgen dem Weg von der Aufnahmeanfrage über die Arbeit mit KI-Entwürfen bis zur wirtschaftlichen Bewertung. Am Ende stehen zwei ausgearbeitete Möglichkeiten, den Nutzen in der eigenen Klinik zu beurteilen: eine geplante wissenschaftliche Untersuchung und eine weniger aufwendige Bewertung mit vorhandenen Daten und ausdrücklich ausgewiesenen Schätzungen. Zwei Arbeitsblätter und eine gesonderte Rechenvorlage helfen bei der Umsetzung.

Die nächste Seite fasst sechs zentrale Überlegungen zusammen. Das Whitepaper verbindet ausgewählte Primärquellen mit eigenen Analysen und Erfahrungen; es ist keine systematische Übersichtsarbeit. Literatur- und Rechtsstand: 13. September 2026.

FÜR DEN SCHNELLEN ÜBERBLICK

Sechs Überlegungen für die Entscheidung über KI

01 Kürzere Antwortzeiten können auf einer neuen Zählweise beruhen

Mit der KI-Einführung kann sich auch der Umgang mit offenen Aufnahmeanfragen ändern. Wird eine Anfrage etwa wegen fehlender Unterlagen geschlossen und später neu angelegt, beginnt die Zeitmessung möglicherweise von vorn. Die ausgewiesene Bearbeitungszeit sinkt dann, obwohl die Person insgesamt ebenso lange warten kann. Der Verlauf ab der ersten Anfrage macht dies sichtbar.

02 Ein gut lesbarer KI-Entwurf kann wichtige Angaben auslassen

Forschende ließen eine KI aus klinischen Notizen zu 100 Krankenhausaufenthalten den Verlauf für den Entlassungsbrief zusammenfassen. Gegenüber den ärztlichen Fassungen bewerteten Prüfende die KI-Texte als knapper und zusammenhängender. Zugleich fehlten mehr Angaben. Gute Lesbarkeit bedeutete hier keine größere Vollständigkeit.

03 Mehrere automatische Kontrollen können dieselbe Lücke übersehen

Geht beim Einlesen einer Akte eine Seite verloren, können Texterstellung und automatische Gegenprüfung dieselbe unvollständige Grundlage verwenden. Ihre Übereinstimmung sagt dann wenig über die Lücke. Eine zusätzliche Prüfung braucht Zugang zu den ursprünglichen Unterlagen oder einen anderen Hinweis auf die fehlende Seite.

04 Schneller erstellte Entwürfe können die Freigabe stärker beanspruchen

Erstellen mehrere Stationsärzt:innen ihre Berichte früher, können bei den freigebenden Oberärzt:innen mehr Entwürfe gleichzeitig eintreffen. Die Schreibzeit sinkt dann möglicherweise, während die Berichte länger auf Freigabe warten. Der Nutzen zeigt sich auch daran, wann der fertige Bericht vorliegt und wie viel Arbeit insgesamt nötig war.

05 Eine Anweisung an die KI kann viele Berichte zugleich verändern

Soll die KI für den Entlassungsentwurf stets das neueste Dokument bevorzugen, klingt das zunächst plausibel. Ein neuerer Brief kann jedoch eine ältere, weiterhin gültige Festlegung ausgelassen haben. Wird dieselbe Anweisung für viele Fälle verwendet, wiederholt sich möglicherweise derselbe Fehler. Eine erprobte Korrektur der Vorlage kann entsprechend vielen Berichten zugutekommen.

06 Der Aufwand der Bewertung hängt von der anstehenden Entscheidung ab

Für eine begrenzte Fortführung in einem Bereich können vorhandene Fallzahlen und gekennzeichnete Zeitschätzungen eine wirtschaftliche Entscheidung unterstützen. Wer belegen will, dass die KI eine Verbesserung verursacht hat, braucht einen geeigneten Vergleich und eine dafür geplante Untersuchung. Beide Bewertungswege werden hier mit ihren jeweiligen Aussagegrenzen ausgearbeitet.

Die Kapitel erläutern die Studien und führen die Beispiele weiter. Die Arbeitsblätter am Ende unterstützen die Vorbereitung einer eigenen Bewertung.

ÜBER DEN HERAUSGEBER

Was aiomics für Kliniken entwickelt

Aiomics entwickelt Software, die klinische Unterlagen zusammenführt, bei der Erstellung medizinischer Dokumente unterstützt und den Bearbeitungsstand von Aufnahmeanfragen für das Team sichtbar macht. Ärzt:innen können beispielsweise Vorbefunde und eigene Untersuchungsnotizen für einen Aufnahmebericht aufbereiten lassen. Für Entlassungsberichte und weitere fallbezogene Schreiben stehen ebenfalls Vorlagen zur Verfügung.

Gemeinsam mit den klinischen Teams passen wir die Vorlagen an Fachgebiete, Berufsgruppen, Kostenträger und örtliche Vorgaben an. Quellenbelege und automatische Prüfungen helfen dabei, erzeugte Texte mit den Ausgangsunterlagen abzugleichen und Unstimmigkeiten zu erkennen. Die behandelnden Fachpersonen treffen die medizinischen Entscheidungen und geben die Ergebnisse frei.

Das Unternehmen wird ärztlich geführt. Dr. med. Sven Jungmann verantwortet die medizinische Perspektive; Dr. Nikita Tarasov bringt als promovierter Physiker die technische und wissenschaftliche Perspektive ein. Wir arbeiten mit klinischen Teams in Deutschland und im Ausland zusammen.

Dieses Whitepaper verbindet veröffentlichte Forschung mit eigenen Analysen und Erfahrungen aus dieser Arbeit. Die gesonderte Unternehmensseite beschreibt unser Angebot. Die fachlichen Überlegungen und Bewertungsblätter lassen sich unabhängig von einer bestimmten Anwendung nutzen.



Dr. med. Sven Jungmann
Mitgründer von aiomics und Autor dieses Whitepapers

Gern gehen wir mit Ihnen einen Ablauf Ihrer Klinik durch und zeigen anhand eines passenden Beispiels, welche Unterstützung möglich ist.

Eine persönliche Vorstellung vereinbaren

Inhalt

	Was sich mit KI in der Rehabilitation verändert	2
	Sechs Überlegungen für die Entscheidung über KI	3
	Was aiomics macht	4
01	Fallauswahl und Belegung	6
02	Arbeit, Engpässe und Zuständigkeit	7
03	Vertrauen in klinische Texte	9
04	Die Grenzen automatischer Prüfung	10
05	Vorlagen und institutionelles Wissen	12
06	Gemeinsame Informationen und aktuelle Dokumente	13
07	Lernen aus Korrekturen und Rückmeldungen	15
	Aiomics in der Rehabilitation	17
08	Wirtschaftlicher Nutzen und Annahmen	18
09	Zwei Wege zur Nutzenbewertung	19
10	Erprobung, Anbindung und Weiterentwicklung	22
	Bewertungsblatt 1: Entscheidung und Datenlage	23
	Bewertungsblatt 2: Unsicherheit und Beschluss	24
	Literatur und Quellen	25

1. Wie KI im Aufnahmemanagement die Fallauswahl verändern kann

Wenn eine KI Vorbefunde und Entlassbriefe für das Aufnahmeteam zusammenführt, kann die Vorbereitung einer Aufnahmeentscheidung weniger Zeit beanspruchen. Was die Klinik dadurch gewinnt, hängt auch davon ab, welche Anfragen sie anschließend anders bearbeitet. Sie könnte geeigneten Patient:innen früher zusagen, schwierige Anfragen gründlicher prüfen oder ungeeignete Fälle früher ablehnen. Alle drei Entwicklungen können die gemessene Antwortzeit verkürzen. Für Belegung und Versorgung bedeuten sie Unterschiedliches.

Kürzere Wartezeiten können mit einer anderen Zählweise einhergehen

Neferu und Fleck untersuchten ein kanadisches Rehabilitationsprogramm für Menschen mit erworbener Hirnschädigung. Anmeldungen erreichten die Aufnahmekoordination nach einer organisatorischen Änderung bereits am Eingangstag statt erst bei der wöchentlichen Besprechung. KI wurde nicht eingesetzt. Im Vergleich der Zeiträume davor und danach sank die mittlere Zeit bis zur Entscheidung von 9,5 auf 4,0 Tage, bis zur Aufnahme von 26,7 auf 12,2 Tage. Gleichzeitig sank die Aufnahmequote von 80 auf 43 Prozent. Nach der Änderung lagen nur zehn Aufnahmen vor; Verbesserungen hatten teilweise schon vorher eingesetzt.¹

Die Autorinnen nennen eine mögliche Erklärung: Anmeldungen noch nicht rehafähiger Personen wurden teilweise abgelehnt, statt offengehalten. Eine spätere erneute Anmeldung zählte als neuer Vorgang. Damit konnten sich der Beginn der gemessenen Wartezeit und der Nenner der Aufnahmequote verändern. Welchen Anteil dieser Umgang mit Anmeldungen an den Ergebnissen hatte, konnte die Untersuchung nicht bestimmen.

Die Aufbereitung der Unterlagen kann beeinflussen, welche Fälle Aufmerksamkeit erhalten

Bei einer Aufnahmeentscheidung können neben der medizinischen Eignung der erwartete Betreuungsaufwand, die personelle Besetzung und Erfahrungen mit der zuweisenden Stelle ins Gewicht fallen. Gerade bei Grenzfällen lässt sich aus der dokumentierten Begründung kaum erkennen, wie stark jede dieser Erwägungen gewirkt hat. Für die Einführung einer KI-Anwendung ist deshalb relevant, ob ihre Aufbereitung der Unterlagen die Abwägung verän-

dert. Das lässt sich beispielsweise daran untersuchen, welche zuvor ungeklärten Anfragen das Aufnahmeteam nun entscheiden kann.

Denkbar sind zwei unterschiedliche Entwicklungen. Bei gut lesbaren Vorbefunden erzeugt die KI rasch eine brauchbare Übersicht; unvollständige oder schlecht lesbare Unterlagen erfordern weiterhin viel Nacharbeit. Bevorzugt das Team die schnell vorbereiteten Anfragen, kann die Qualität der Dokumentation mittelbar die Fallauswahl beeinflussen. Eine Anwendung, die auch schwer zugängliche Vorbefunde zuverlässig erschließt, könnte dagegen den Bearbeitungsaufwand gerade bei komplexen Anfragen senken. Dies sind mögliche Folgen, die bei der örtlichen Einführung zu prüfen wären. Aus einer kürzeren durchschnittlichen Antwortzeit lässt sich keine von beiden ablesen.

Auch die Weiterleitung einer Anfrage kann die eigene Arbeitsliste verkürzen, während die Klärung für die betroffene Person offenbleibt. Übernimmt die nächste Einrichtung tatsächlich, oder beginnt dort dieselbe Suche nach Zuständigkeit und Unterlagen erneut? Wer den Nutzen einer gemeinsamen Fallübersicht beurteilt, braucht daher einen Blick auf den weiteren Verlauf weitergegebener Anfragen. Die Kennzeichnung „abgeschlossen“ im eigenen Bereich beschreibt zunächst nur, dass dieser Bereich seine Bearbeitung beendet hat.

Kosten, Fallauswahl und Behandlungsintensität getrennt betrachten

Sood und Kolleg:innen untersuchten, wie US-Rehabilitationseinrichtungen auf eine Änderung ihrer Vergütung reagierten. Sie werteten Daten des US-Versicherungsprogramms Medicare von 2001 bis 2003 zu Patient:innen nach Schlaganfall, Hüftfraktur oder Gelenkersatz an der unteren Extremität aus. Ihre Analyse unterschied Änderungen der Aufnahmeentscheidungen, der Zusammensetzung der aufgenommenen Fälle und der Behandlung. Die nach der Vergütungsumstellung zunächst sinkenden Kosten führten sie vor allem auf eine veränderte Behandlungsintensität zurück; zugleich wurde die Aufnahme ausgeweitet. Die Ergebnisse beschreiben eine ältere US-Vergütungsreform. Für eine deutsche Rehaklinik ist vor allem der Untersuchungsansatz nützlich: Sinkende Kosten je Fall lassen sich erst einordnen, wenn auch Veränderungen der aufgenommenen Patientengruppe und der erbrachten Behandlung betrachtet werden.²

Für eine KI-gestützte Aufbereitung von Aufnahmeanfragen ergibt sich daraus eine konkrete Frage:

Welche bisher unsicheren Entscheidungen würden anders ausfallen, wenn Vorbefunde, funktionelle Einschränkungen und erwarteter Unterstützungsbedarf vollständig und verlässlich vorlägen? Bleibt die Anfrage auch dann ungeeignet, liegt die Grenze eher bei den Versorgungsmöglichkeiten der Klinik. Wird sie durch die bessere Übersicht entscheidbar, könnte die Anwendung bisherige Unsicherheit verringern. Eine Besprechung solcher Fälle hilft, den erhofften Nutzen vor der Einführung genauer zu benennen.

In der späteren Auswertung sollten deshalb die eingegangenen Anfragen vor und nach der Einführung verglichen werden. Werden medizinisch und organisatorisch ähnlich anspruchsvolle Fälle schneller entschieden? Verschieben sich die Anteile der Fallgruppen? Ändert sich das Verhältnis von Zusagen, Absagen und Nachforderungen? Erneute Anmeldungen derselben Person bleiben dabei erkennbar. Erst mit diesem Zusammenhang wird verständlich, ob eine veränderte Aufnahmequote auf andere Anfragen, andere Entscheidungen oder eine andere Zählweise zurückgeht.

Bei Leistungen der Deutschen Rentenversicherung hängt die Zuweisung zudem von der trägerseitigen Einrichtungsauswahl und dem Wunsch- und Wahlrecht ab. Die Deutsche Rentenversicherung nennt unter anderem Indikation, besondere Anforderungen, Qualität, Entfernung und Wartezeit als Auswahlkriterien. Eine schnellere oder verlässlichere Rückmeldung der Klinik ist somit nur ein Teil der Bedingungen, unter denen eine Aufnahme zustande kommt. Eine daraus abgeleitete Belegungssteigerung müsste die Klinik gesondert nachweisen.³

Eine aussagekräftige Nutzenprüfung schließt Anfragen ein, die bisher viel Klärung erfordert haben. Werden auch diese leichter entscheidbar, reicht der Nutzen der KI-Aufbereitung weiter als die Beschleunigung ohnehin gut dokumentierter Fälle.

2. Wie sich Schreibaarbeit, Freigaben und Abstimmung neu verteilen

Angenommen, mehrere Stationsärzt:innen können ihre Entlassungsentwürfe mit KI früher fertigstellen. Für die fachliche Freigabe stehen weiterhin dieselben Oberärzt:innen zur Verfügung. Treffen nun mehr prüfbereite Entwürfe gleichzeitig ein, kann vor deren Freigabe eine neue Warteschlange entstehen. Die Stationsärzt:innen haben weniger Schreibarbeit, während die Berichte möglicherweise kaum

früher abgeschlossen werden. Dieses Beispiel zeigt, weshalb eingesparte Schreibzeit und Zeit bis zur Freigabe gemeinsam betrachtet werden sollten.

Die Leitung kann aus diesem Befund unterschiedliche Konsequenzen ziehen. Vielleicht müssen Freigaben anders über den Tag verteilt werden. Vielleicht erschweren lange, ungegliederte Entwürfe die Gegenlesung. Vielleicht wächst durch die Unterstützung die Zahl bearbeitbarer Fälle, sodass zusätzliche fachliche Prüfzeit benötigt wird. Welche Erklärung zutrifft, zeigt sich erst, wenn die Arbeit zwischen Entwurf und Freigabe mit erfasst wird. Die Summe eingesparter Minuten bei der Texterstellung reicht dafür nicht aus.

Eine Fallübersicht hilft, wenn das Team offene Aufgaben damit früher klärt

Rolls und Kolleg:innen untersuchten ein australisches Rehabilitationszentrum, das eine elektronische Übersicht zum Verlauf der Patient:innen eingeführt hatte. Sie verglichen rückblickend 3.259 Rehabilitationsepisoden vor und nach der Einführung. Nach statistischer Berücksichtigung von Unterschieden zwischen den Fällen war die Verweildauer um 4,1 Tage kürzer. Ohne gleichzeitige Vergleichseinrichtung lässt sich der Anteil der elektronischen Übersicht an dieser Veränderung jedoch nicht sicher bestimmen.⁴ Eine britische Untersuchung von Mebrahtu und Kolleg:innen betrachtete die Einführung eines zentralen Krankenhaus-Leitzentrums mit elektronischen Übersichten und begleitenden organisatorischen Änderungen. Sie fand über die untersuchten Phasen hinweg keine durchgehend verlässliche Verbesserung der Aufnahme-, Behandlungs- und Verlegungsabläufe. Auch hier erschwerten gleichzeitige Veränderungen im Betrieb die Zuschreibung. Die beiden Studien untersuchten unterschiedliche Einrichtungen und Maßnahmen; sie sind kein direkter Vergleich zweier Produkte.⁵

Für eine Rehaklinik wäre deshalb zu klären, welche Entscheidung durch eine neue Übersicht früher möglich werden soll. Soll das Aufnahmeteam sehen, wer eine fehlende Voruntersuchung nachfordert? Soll die ärztliche Leitung erkennen, welche Entwürfe auf eine Freigabe warten? Eine Tafel nach dem Kanban-Prinzip stellt Fälle als Karten in Spalten mit benannten Bearbeitungsständen dar. Ihr organisatorischer Nutzen hängt daran, dass die Spalten eine gemeinsame Bedeutung haben und ein Wechsel des Bearbeitungsstands die vereinbarte nächste Handlung auslöst.

Eine gemeinsame Übersicht macht Zuständigkeitsgrenzen verhandelbar

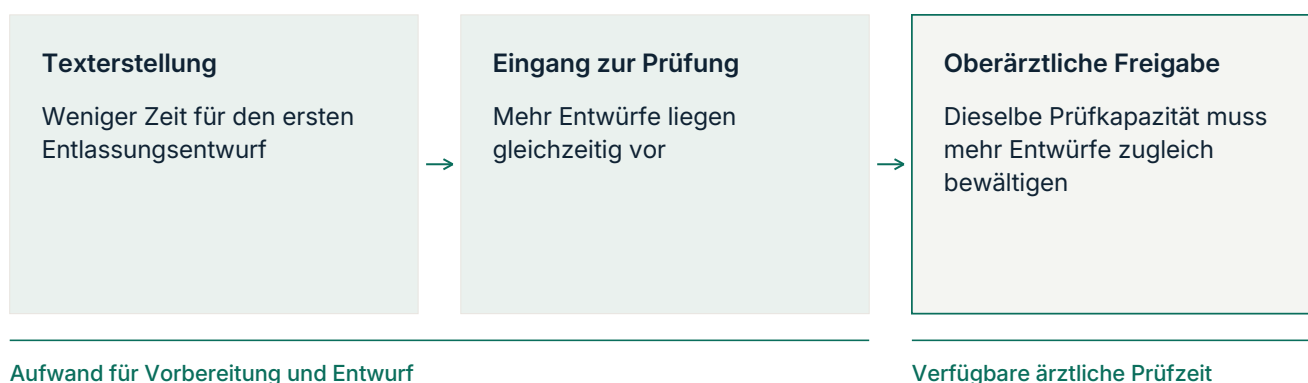
Wird eine Aufnahmeanfrage telefonisch oder per E-Mail weitergegeben, können abgebende und empfangende Stelle unterschiedlich beurteilen, wann die Aufgabe übernommen wurde. Eine gemeinsame Fallübersicht kann diesen Übergang sichtbar machen: Wer hat die Anfrage erhalten, wer bearbeitet sie, welche Klärung steht aus? Damit werden zugleich bisherige Handlungsspielräume berührt. Wenn Teams über die Bedeutung eines Bearbeitungsstands streiten, kann dahinter eine ungeklärte Zuständigkeit liegen. Eine zusätzliche Schulung zur Bedienung löst diesen Konflikt allein kaum.

Bei der Kennzahl Antwortzeit wird die Bedeutung solcher Vereinbarungen besonders deutlich. Beginnt die Messung erst, nachdem das Aufnahmeteam sämtliche Unterlagen als vollständig markiert hat, fehlt die Wartezeit auf diese Unterlagen. Endet sie bereits mit einer Nachforderung, misst sie die erste Reaktion und noch keine Aufnahmeentscheidung. Beide Größen können nützlich sein, beantworten aber verschiedene Fragen. Werden sie vor und nach einer Einführung unterschiedlich erhoben, entsteht leicht der Eindruck einer Beschleunigung, obwohl Patient:innen weiterhin ebenso lange auf eine Entscheidung warten.

SCHAUBILD 1

Schneller geschriebene Berichte können länger auf Freigabe warten

Ein mögliches Beispiel für die Verlagerung eines Engpasses; keine gemessene Wirkung.



Gesamter Ablauf: einschließlich Prüfung, Korrektur und Rückfragen.

In diesem Beispiel sinkt die Schreibzeit je Bericht. Wenn sich die Freigabe verzögert, kann der fertige Bericht trotzdem später vorliegen. Für die Bewertung werden beide Zeiten benötigt.

Eine örtliche Nutzenrechnung sollte daher auch die Arbeit erfassen, die andere Stellen zusätzlich übernehmen. Mehr Nachforderungen können das Aufnahmeteam entlasten und die zuweisende Klinik stärker beanspruchen. Ein einfacher Datenimport kann wiederholte Eingaben im eigenen Bereich vermeiden, aber neue Zuordnungsarbeit im Nachbarbereich erzeugen. Für die Leitung eines Klinikverbunds ist entscheidend, ob über die beteiligten Bereiche hinweg weniger Arbeit anfällt oder ob sie lediglich an anderer Stelle geleistet wird.

Entfallende Statusfragen und notwendige Rückfragen unterscheiden

Weigl und Kolleg:innen begleiteten 29 Krankenhausärzt:innen während 43 Diensten und erfassten Unterbrechungen sowie die wahrgenommene Arbeitsbelastung. Häufigere Unterbrechungen waren mit höherer Belastung verbunden. Die Beobachtungsstudie untersuchte weder KI noch elektronische Fallübersichten; sie bestimmte auch nicht, welche einzelnen Unterbrechungen klinisch entbehrlich waren.⁶ Für eine Einführung lässt sich daraus eine präzisere Frage ableiten: Welche Kontakte dienen lediglich dazu, den Stand eines Falls zu

erfahren, und welche bringen eine notwendige fachliche Klärung voran?

Eine verlässliche Übersicht könnte beispielsweise den Anruf ersetzen, ob eine Anfrage bereits geprüft wird. Macht dieselbe Übersicht eine widersprüchliche Einschätzung sichtbar, kann sie zugleich eine zusätzliche fachliche Rückfrage auslösen. Weniger Nachrichten sind deshalb kein hinreichendes Erfolgskriterium. Aussagekräftiger wäre, ob Such- und Statusarbeit abnimmt und die benötigte Expertise weiterhin rechtzeitig erreichbar ist.

3. Wie sich die Verlässlichkeit KI-erstellter Berichtsentwürfe beurteilen lässt

Ein KI-erstellter Entlassungsentwurf kann sprachlich geschlossen wirken, obwohl die zugrunde liegende Akte widersprüchlich oder unvollständig ist. Für die empfangende Rehaklinik entsteht dann eine besondere Schwierigkeit: Brüche, die beim Lesen sonst Anlass zum Nachfragen geben würden, können in der Zusammenfassung verschwunden sein. Ob diese Glättung auch die fachliche Darstellung verbessert hat, lässt sich aus dem Stil allein nicht erkennen. Studien zu KI-erstellten klinischen Texten helfen, die verschiedenen Fehlerquellen auseinanderzuhalten.

Ein verständlicher Bericht kann entscheidende Angaben auslassen

Williams und Kolleg:innen untersuchten 100 Krankenhausaufenthalte in den USA. GPT-4 Turbo sollte aus den klinischen Notizen jeweils den erzählenden Teil des Entlassungsbriefts entwerfen, der Aufnahme und stationären Verlauf beschreibt. Ärztliche Prüfende verglichen diese Texte mit den ärztlich verfassten Abschnitten derselben Aufenthalte und glichen beide Fassungen mit den Notizen ab. Sie wussten nicht, welche Fassung von der KI stammte. Die KI-Texte waren knapper und zusammenhängender, aber weniger vollständig. Im Mittel enthielten sie 2,91 unterschiedliche Fehler gegenüber 1,82 in den ärztlichen Fassungen, darunter häufiger Auslassungen und unrichtige Angaben. Gesamtqualität und Präferenz unterschieden sich nicht statistisch signifikant.⁷

Der Befund erklärt eine Schwierigkeit der Gegenlesung: Beim Lesen beurteilt man zunächst den vorliegenden Text. Was darin fehlt, wird oft erst beim erneuten Blick in die Akte erkennbar. Eine Dokumentationshilfe kann diese zweite Prüfrichtung

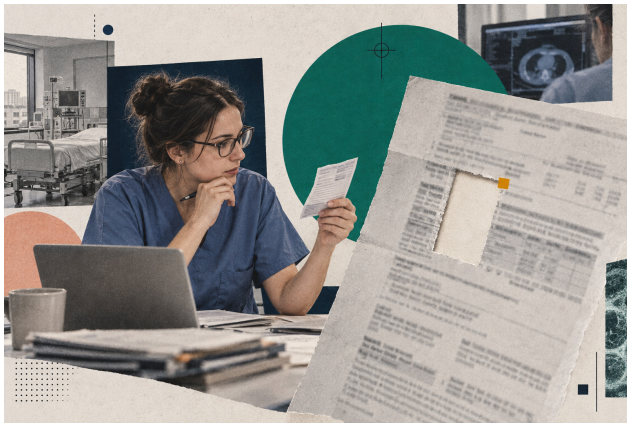
unterstützen, indem sie die maßgeblichen Ausgangsunterlagen unmittelbar zugänglich macht.

Bei diesem Abgleich kommt es auch darauf an, wann eine Aussage galt. Ein Aufnahmebefund kann einen Verdacht enthalten, der im Verlauf verworfen wurde. Wenn eine KI beide Einträge zu einer einheitlichen Erklärung zusammenzieht, kann der Unterschied zwischen anfänglicher Vermutung und abschließender Beurteilung verloren gehen. Ein Quellenverweis hilft dann nur, wenn die prüfende Person auch die zeitliche Einordnung und die spätere Änderung nachvollziehen kann. Die sprachlich stimmige Darstellung ist andernfalls erst durch das Zusammenfassen entstanden.

Auch eine belegte Aussage kann einen Fehler aus der Akte fortschreiben

Omar und Kolleg:innen prüften sechs Sprachmodelle an 300 ärztlich konstruierten Fallbeschreibungen. In jede Beschreibung hatten sie gezielt eine erfundene Untersuchung, ein erfundenes klinisches Zeichen oder eine erfundene Erkrankung eingefügt. Die Modelle sollten die Fälle bearbeiten und übernehmen die falschen Vorgaben häufig; teilweise entwickelten sie dazu weitere Erklärungen. Dieser gezielte Belastungsversuch beschreibt keine Fehlerhäufigkeit im klinischen Alltag. Er zeigt, wie aus einem falschen Ausgangspunkt eine ausführlicher begründete, weiterhin falsche Antwort entstehen kann.⁸

Für die Erstellung eines Reha-Berichts ist dieser Mechanismus relevant, wenn etwa ein früherer Arztbrief bereits eine unzutreffende Angabe enthält. Eine KI kann die Angabe korrekt übernehmen und mit einem passenden Quellenverweis versehen. Damit ist ihre Herkunft nachvollziehbar. Ob der Inhalt zutrifft, zum aktuellen Zustand passt oder in einem späteren Dokument korrigiert wurde, muss dennoch geklärt werden. Der Quellenbeleg erleichtert diese Prüfung, beantwortet sie aber nicht von selbst.



Der Quellenverweis zeigt, aus welchem Dokument eine Aussage stammt. Beim Abgleich bleibt zu klären, ob sie zutrifft, wann sie galt und ob spätere Befunde sie verändert haben.

Vollständig fehlende Angaben verlangen einen zweiten Blick auf die Ausgangsakte

Die Forschungsarbeit NOHARM untersuchte Antworten von KI-Systemen auf medizinische Konsultationsfragen. Die Aufgaben beruhten auf Anfragen der Primärversorgung an zehn Fachgebiete; die Systeme sollten Vorschläge für die weitere Versorgung formulieren. Fachleute bewerteten, welche Fehler oder fehlenden Schritte bei einer Umsetzung der Antworten schaden könnten. In der hier verwendeten Vorabfassung entfielen mehr als 80 Prozent der als schwerwiegend eingestuften Fehler auf notwendige Schritte, die in den Antworten fehlten. Tatsächliche Patientenschäden wurden nicht gemessen. Da Versorgungsempfehlungen untersucht wurden, lässt sich daraus keine Fehlerquote für Reha-Berichte ableiten.⁹

Für die Kontrolle eines Berichtsentwurfs ergibt sich daraus eine ergänzende Prüfrichtung. Wer jeden Satz des Entwurfs mit einer Quelle abgleicht, findet möglicherweise alle unzutreffenden Aussagen und übersieht trotzdem eine wichtige Angabe, die im Entwurf überhaupt nicht vorkommt. Deshalb muss die Kontrolle auch bei den maßgeblichen Befunden und Festlegungen der Akte ansetzen: Sind sie im Entwurf enthalten und mit ihrer Bedeutung erhalten geblieben? Ein lückenlos belegter Text kann in diesem Sinn weiterhin unvollständig sein.

Für die ärztliche Gegenlesung verändert sich damit die Verteilung der Aufmerksamkeit. Soweit die KI Orthografie und Aufbau zuverlässig übernimmt, kann mehr Prüfzeit für zeitliche Widersprüche, fehlende Festlegungen und unzutreffende Schlussfol-

gerungen verfügbar werden. Ob die Anwendung insgesamt entlastet, hängt auch davon ab, wie leicht Ärzt:innen die betreffenden Quellen öffnen und mit dem Entwurf vergleichen können. Ein gut formulierter Text ohne solchen Zugang kann weiterhin erheblichen Suchaufwand verursachen.

4. Was automatische Kontrollen eines KI-Entwurfs erkennen können

Eine Dokumentationsanwendung kann einen Entwurf erzeugen und anschließend ein weiteres Modell mit dessen Kontrolle beauftragen. Das zweite Modell könnte beispielsweise prüfen, ob die Aussagen im Entwurf durch die eingelesenen Unterlagen gedeckt sind. Wie viel zusätzliche Absicherung entsteht, hängt davon ab, welche Ausgangsdaten diese Kontrolle erhält. Greift sie auf dieselbe fehlerhaft erfasste Fassung zurück, kann der ursprüngliche Fehler beide Schritte passieren. Für die Beschaffung ist deshalb die Art des Abgleichs ebenso wichtig wie die Zahl der Kontrollen.

Mehrere Kontrollen können denselben Erfassungsfehler übernehmen

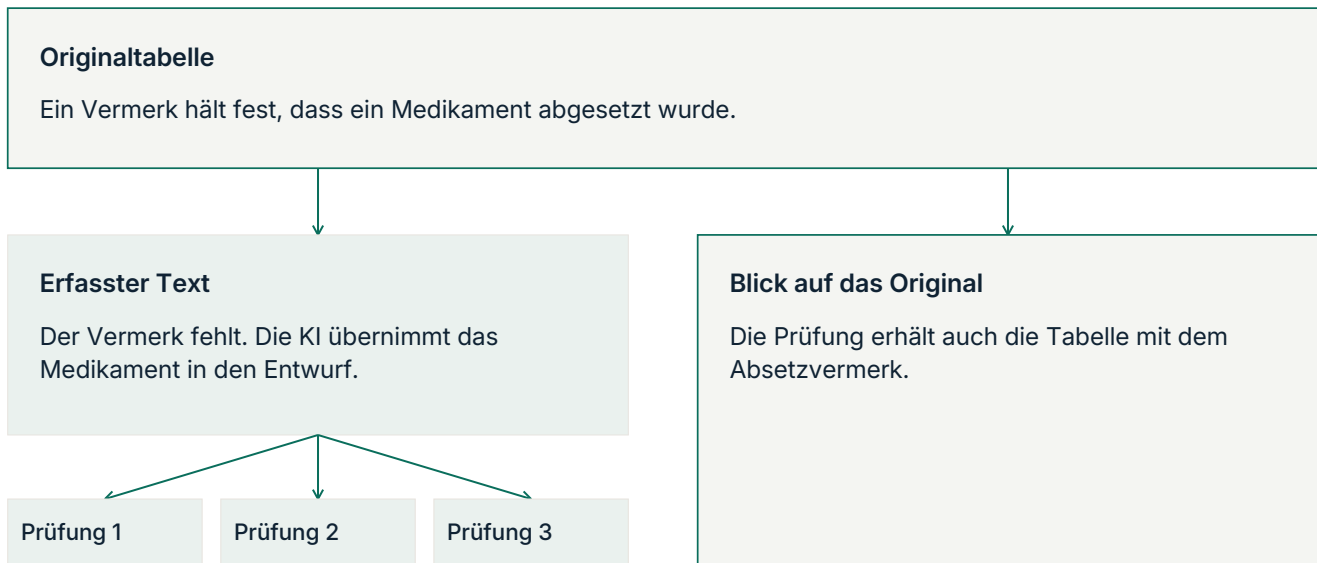
Ein vereinfachtes Beispiel: Beim Einlesen einer schlecht lesbaren Labortabelle wird ein Wert dem falschen Untersuchungsdatum zugeordnet. Der Berichtsentwurf übernimmt diese Zuordnung. Drei nachfolgende Kontrollen vergleichen den Entwurf mit der bereits erfassten Tabelle und finden keinen Widerspruch. Alle bestätigen damit eine Übereinstimmung, obwohl das Datum im Original anders lautet. Ein zusätzlicher Abgleich mit der Originaltabelle eröffnet die Möglichkeit, gerade diesen Erfassungsfehler zu entdecken. Ob der Abgleich ihn tatsächlich erkennt, müsste am Beispiel geprüft werden.

Dasselbe Problem entsteht, wenn eine Seite bei der Erfassung fehlt oder die Dokumentensuche einen späteren Korrekturbeitrag nicht auswählt. Eine Kontrolle, die nur die ausgewählten Textstellen erhält, kann die nicht vorgelegte Seite oder Korrektur nicht berücksichtigen. Mehrere übereinstimmende Modellantworten belegen daher für sich genommen keine voneinander unabhängige Absicherung. Diese Beispiele erläutern mögliche Fehlerwege; sie sind kein klinischer Sicherheitsnachweis für eine bestimmte technische Lösung.

SCHAUBILD 2

Ein Abgleich mit dem Original kann einen Erfassungsfehler aufdecken

Vereinfachtes Beispiel: Beim Einlesen wird eine Angabe aus einer Medikamententabelle ausgelassen.



Alle vergleichen mit dem erfassten Text und finden keinen Widerspruch.

Die drei Kontrollen bestätigen die Übereinstimmung mit einer unvollständig erfassten Quelle. Der zusätzliche Zugriff auf die Originaltabelle macht den fehlenden Vermerk prüfbar. Ob die Kontrolle ihn erkennt, muss erprobt werden.

Vom Eingangsdokument bis zum freigegebenen Text getrennt prüfen

Die Prüfung kann bereits beim Eingang der Unterlagen beginnen: Sind alle vorgesehenen Dokumente und Seiten vorhanden? Beim Erfassen ist zu klären, ob Werte, Datumsangaben und Zuordnungen aus dem Original erhalten bleiben. Beim Zusammenfassen geht es zusätzlich darum, ob etwa eine spätere Korrektur als solche erkennbar bleibt. Erst danach lässt sich beurteilen, welche Aussagen der Entwurf enthält und welche medizinischen Fragen weiterhin offen sind. Eine fachlich plausible Zusammenfassung beantwortet die vorgelagerten Fragen zur Vollständigkeit und Erfassung ihrer Quellen noch nicht.

Grolleau und Kolleg:innen untersuchten die Nutzung KI-erstellter Zusammenfassungen des stationären Verlaufs in einem US-Krankenhaus. Die Anwendung arbeitete mit wiederholter Texterstellung und Quellenabgleich; elf Ärzt:innen erprobten sie über zehn Wochen bei insgesamt 384 Entlassungen. Zu 100 Entwürfen lagen ärztliche Rückmeldungen zur Sicherheit vor. In 25 wurden Auslassungen, in 20 unrichtige Angaben und in zwei erfundene Inhalte gemeldet; mehrere Fehlerarten konn-

ten im selben Entwurf auftreten. Nur ein Teil der Entwürfe wurde auf diese Weise beurteilt, und eine gleichzeitige Vergleichsgruppe fehlte. Die Zahlen beschreiben deshalb die Rückmeldungen zu diesen 100 Entwürfen. Sie zeigen, dass auch nach mehreren automatischen Bearbeitungsschritten noch Fehler von Ärzt:innen gefunden wurden.¹⁰

Auch Entwürfe ohne Warnung in die Qualitätsprüfung einbeziehen

Die Zahl automatischer Warnungen ist ohne weitere Angaben schwer zu deuten. Viele gemeldete Widersprüche können auf eine wirksame Fehlersuche, auf besonders widersprüchliche Unterlagen oder auf zahlreiche belanglose Meldungen zurückgehen. Ebenso mehrdeutig ist die Zahl ärztlicher Korrekturen: Wenige Änderungen können gute Entwürfe anzeigen, aber auch übersehene Fehler. Zur Qualitätsbeurteilung gehört deshalb eine fachliche Prüfung ausgewählter Entwürfe unabhängig davon, ob das System sie beanstandet hat.

Eine Stichprobe sollte entsprechend auch Entwürfe ohne automatische Warnung enthalten. Bei ihnen lässt sich untersuchen, welche Fehler die Anwendung nicht gemeldet hat. Wählt eine Klinik für die

vertiefte Gegenlesung absichtlich besonders viele auffällige Entwürfe aus, kann sie damit gezielt Schwachstellen suchen. Für eine Fehlerquote über alle Entwürfe muss sie jedoch die unterschiedlichen Auswahlwahrscheinlichkeiten berücksichtigen. Sonst beschreibt die Quote vor allem die Zusammensetzung der vertieft geprüften Fälle.

Für einen Anbietervergleich lassen sich solche Gegenlesungen an gezielt schwierigen, datenschutzgerecht vorbereiteten Prüffällen durchführen. Bei einem Fall fehlen Seiten, bei einem zweiten widersprechen sich frühe und späte Befunde, bei einem dritten ist eine Tabelle schlecht lesbar. Die Klinik kann beobachten, ob die Anwendung ihre Grenzen erkennbar macht oder trotzdem eine geschlossene Darstellung erzeugt. Werden dieselben Fälle nach einem Modell- oder Vorlagenwechsel erneut bearbeitet, wird auch sichtbar, ob eine früher behobene Schwachstelle zurückkehrt.

Der Freigabeschritt muss schließlich zur Arbeit der prüfenden Ärzt:innen passen. Können sie die maßgeblichen Originalbefunde unmittelbar aufrufen? Steht für komplexe Fälle ausreichend Gegenlesezeit zur Verfügung? Ist erkennbar, welche Unstimmigkeiten noch ungeklärt sind? Eine vorhandene Freigabeschaltfläche sagt darüber wenig aus. Erst die tatsächlichen Arbeitsbedingungen zeigen, welche fachliche Kontrolle zwischen Entwurf und freigegebenem Bericht stattfinden kann.

5. Wie gemeinsame Prompts die Begründung in Reha-Berichten prägen

Die Anweisungen an die KI werden in einem Prompt zusammengefasst. Nutzt eine Klinik ihn für viele Entlassungsberichte, gelten seine Regeln bei jeder Bearbeitung. Der Prompt kann etwa verlangen, spätere fachliche Korrekturen ausdrücklich zu berücksichtigen oder widersprechende Einschätzungen verschiedener Berufsgruppen nebeneinander sichtbar zu lassen. Er kann aber auch pauschal die jeweils jüngste Quelle bevorzugen und dadurch eine wichtige frühere Festlegung verdrängen. Die fachliche Reichweite einer gemeinsamen Vorlage hängt deshalb an einzelnen Vorgaben, die beim Lesen eines gelungenen Musterberichts leicht unauffällig bleiben.

Einheitliche Gliederung und individuelle Begründung gemeinsam erhalten

Der DRV-Leitfaden für den ärztlichen Reha-Entlassungsbericht gibt eine gemeinsame Gliederung vor. Innerhalb dieser Gliederung soll die sozialmedizinische Beurteilung aus den Befunden, der Funktionsfähigkeit, den beruflichen Anforderungen und dem Verlauf nachvollziehbar werden.¹¹ Für einen KI-Entwurf bedeutet das: Die Abschnitte können in jedem Bericht gleich heißen, während die Begründung die Besonderheiten der jeweiligen Person wiedergeben muss. Eine automatisch vervollständigte Gliederung lässt noch offen, ob dieser fallbezogene Zusammenhang im Text erkennbar ist.

Die Anwendung könnte beispielsweise ärztliche Befunde, therapeutische Beobachtungen und die dokumentierten beruflichen Anforderungen zusammenführen. Bei zwei Personen mit ähnlichen Diagnosen können sich daraus unterschiedliche Beurteilungen ergeben. Fasst die KI beide Fälle mit nahezu derselben Begründung zusammen, wäre zu prüfen, welche Unterschiede verloren gegangen sind. Entscheidend ist dabei der Bezug zwischen den vorliegenden Befunden und der jeweiligen Schlussfolgerung; sprachliche Abwechslung allein würde den Mangel nicht beheben.

Auch im somatischen Begutachtungsverfahren der DRV dienen anonymisierte Entlassungsberichte und individuelle Therapiepläne als Grundlage einer fachlichen Prüfung nach einem festgelegten Handbuch.¹² Für die örtliche Erprobung eines KI-Entwurfs lässt sich daran eine konkrete Frage anschließen: Kann eine prüfende Fachperson aus dem Bericht nachvollziehen, weshalb gerade bei dieser Person die dokumentierte Leistungsbeurteilung zustande kommt? Ein formal vollständiger Abschnitt kann dafür zu allgemein bleiben.

Vorlagen an ähnlichen Fällen mit unterschiedlichen Folgen erproben

Besonders aufschlussreich sind deshalb Prüffälle, die klinisch ähnlich wirken, sich aber in beruflichen Anforderungen, Alltagsbedingungen oder Unterstützungsbedarf unterscheiden. Die Gegenlesung zeigt dann, ob der Prompt die Begründung auf die jeweilige Person bezieht. Ein zweiter Prüfstein sind widersprechende Einschätzungen im Team oder ein Verlauf, der das ursprüngliche Rehabilitationsziel verändert hat. Hier lässt sich beobachten, ob der Entwurf den Unterschied sichtbar erhält oder ihn beim Zusammenfassen beseitigt.

Bei aiomics gliedern wir Prompts in Zielgruppe, Stil und Aufgabe sowie zusätzliche Vorgaben bei Bedarf; diese Struktur nennen wir „3+1“. Im folgenden Beispiel soll die KI Berichtsentwürfe auf ihre individuelle Begründung prüfen. Dafür erhält sie die jeweiligen Ausgangsunterlagen, die Entwürfe und die bezeichnete Fassung des Prüfhandbuchs. Die Gliederung macht deutlich, an wen die Rückmeldung gerichtet ist, wie sie formuliert werden soll und welchen Vergleich die KI durchführen soll. Die fachliche Beurteilung der Rückmeldung bleibt bei den Prüfenden.

PROMPT 1

Prüfen, ob die Begründung im Entwurf zum einzelnen Fall passt



ZIELGRUPPE. Ärztliche Leitung und die an der Berichtsprüfung beteiligten Fachpersonen. Als Grundlage erhältst du mehrere anonymisierte oder erfundene Prüffälle. Zu jedem Fall gehören Ausgangsunterlagen und ein Berichtsentwurf. Hinzu kommt das ausdrücklich bezeichnete Prüfhandbuch.

STIL. Formuliere präzise und belegbezogen. Halte Selbstauskunft, Beobachtung, Interpretation und abschließende ärztliche Festlegung auseinander.

AUFGABE. Vergleiche jeden Berichtsentwurf mit den Ausgangsunterlagen desselben Falls. Untersuche, ob sich die individuelle Begründung im Entwurf aus diesen Unterlagen nachvollziehen lässt. Benenne allgemeine Formulierungen, die einen fallbezogenen Zusammenhang verdecken, zusammengezogene widersprechende Einschätzungen und nicht belegte Schlussfolgerungen. Gib jeweils die betroffene Aussage und die zugehörige Quelle an. Formuliere eine konkrete Rückfrage an die fachlich Prüfenden; ergänze keine medizinische Beurteilung.

ZUSÄTZLICHE VORGABEN. Vergleiche außerdem die Ausgangsfälle miteinander: Welche Unterschiede sind dort dokumentiert, fehlen aber in den jeweiligen Entwürfen? Verwende ausschließlich die bezeichnete Fassung des Prüfhandbuchs [Fassung, Datum]. Deine Rückmeldung wird anschließend fachlich bewertet.

Änderungen am Prompt an bewährten Prüffällen nachvollziehen

KI kann auch beim Formulieren und Überarbeiten eines Prompts helfen. Dabei sollte die fachlich zuständige Person prüfen, welche Regeln die KI selbst ergänzt hat. Eine scheinbar hilfreiche Anweisung wie „bevorzuge stets die neueste Quelle“ könnte eine spätere beiläufige Angabe über eine ältere ausdrückliche Festlegung stellen. Die Aufforderung, einen widerspruchsfreien Text zu erzeugen, könnte dazu führen, dass die KI fachliche Meinungsunterschiede verdeckt. Solche Ergänzungen verlangen eine bewusste Entscheidung, bevor die Vorlage für weitere Fälle eingesetzt wird.

Eine geänderte Vorlage kann zahlreiche Berichte zugleich verbessern oder dieselbe Fehlannahme in ihnen wiederholen. Deshalb sollten die verantwortlichen Fachpersonen nachvollziehen können, welche Fassung verwendet und was daran geändert wurde. Fälle, bei denen eine frühere Fassung wichtige Unterschiede verloren hat, eignen sich als wiederverwendbare Prüfung: Bleibt die damalige Schwachstelle nach der Überarbeitung behoben?

So wird die Erfahrung aus einzelnen Korrekturen bei der nächsten Änderung des gemeinsamen Prompts erneut nutzbar.

6. Wie Korrekturen an Fallangaben alle betroffenen Teams erreichen

Eine Ärztin lässt aus Vorbefunden, Aufnahmeuntersuchung und eigenen Notizen mehrere Texte vorbereiten: einen ärztlichen Aufnahmebericht, eine Übersicht für die Pflege und die Angaben für Therapieplanung und Sozialdienst. Mit KI wird diese Aufbereitung leichter. Damit vervielfältigen sich auch die Stellen, an denen eine überholte Annahme weiterwirken kann. Entscheidend wird, wie die Klinik eine spätere Korrektur in alle betroffenen Dokumente und Arbeitsabläufe übernimmt.

Eine geänderte Angabe muss alle betroffenen Teams erreichen

Angenommen, im Aufnahmegespräch wurde festgehalten, dass ein Angehöriger zu Hause unterstützen kann. Diese Angabe geht in mehrere der

erzeugten Texte ein. Später stellt sich heraus, dass die Unterstützung ausfällt. Der Sozialdienst braucht nun eine andere Grundlage für die Entlassungsvorbereitung; auch Therapieziele und Hilfsmittelplanung können betroffen sein. Bleibt die Korrektur im ärztlichen Bericht, arbeiten die anderen Teams womöglich mit der alten Angabe weiter.

Bei der Auswahl einer Anwendung lässt sich genau dieser Fall erproben: Welche Texte beruhen noch auf der früheren Angabe? Erhalten die zuständigen Personen einen Hinweis? Ist erkennbar, ob die neue Information bereits fachlich beurteilt und übernommen wurde? Eine gemeinsame Datenablage beantwortet diese Fragen allein noch nicht. Der Anbieter sollte zeigen können, welche Änderungen die Anwendung nachvollzieht und welche Schritte eine verantwortliche Person auslösen muss.

Welche Angaben Pflege, Therapie und Sozialdienst benötigen, lässt sich mit den jeweiligen Empfänger:innen festlegen. Für die Klinikleitung kommt eine weitere Anforderung hinzu: Der Bezug zur gemeinsamen Ausgangsinformation muss erhalten bleiben. Werden die erzeugten Übersichten anschließend als voneinander unabhängige Dokumente fortgeschrieben, muss das Team jede spätere Änderung erneut abgleichen. Dann kann ein Teil der eingesparten Schreibarbeit als Abstimmungsarbeit zurückkehren.

Aufnahmebefund und abschließende Beurteilung zeitlich auseinanderhalten

Ein ähnliches Problem entsteht, wenn die KI einen seit Aufnahme fortgeschriebenen Entlassungsentwurf als Hauptquelle erhält. Darin können eine frühe Verdachtsdiagnose, ein vorläufiges Therapieziel und die spätere Beurteilung nebeneinanderstehen. Das Modell soll daraus einen zusammenhängenden Verlauf formulieren. Fehlt der erkennbare Zeitpunkt einer Aussage, kann es eine überholte Annahme in die abschließende Darstellung übernehmen, obwohl die Verlaufsakte bereits ihre Korrektur enthält.

In unserer Arbeit verwenden wir deshalb eine eigenständige Aufnahmedokumentation als Ausgangspunkt für die beteiligten Berufsgruppen. Befunde, Entscheidungen und Verlauf werden weiterhin zeitnah in der Patientenakte festgehalten. Der erste Entlassungsentwurf wird kurz vor der Entlassung aus diesen Unterlagen und den bis dahin getroffenen Festlegungen vorbereitet; geltende Dokumentations- und Übermittlungsfristen sind dabei zu berücksichtigen. Diese Arbeitsweise soll

die wiederholte Überarbeitung vorläufiger Entlassungstexte verringern. Ihr Nutzen zeigt sich daran, ob weniger Korrekturen und Rückfragen bis zur Freigabe nötig werden.

Der ergänzende Leitfaden „Entlassbriefe mit KI“ auf aiomics.io beschreibt diese Arbeitsweise mit Prompt-Beispielen und einem zusätzlichen Patientenbrief. Für die Organisation bleibt vor allem festzulegen, welche Dokumente den Aufnahmezustand festhalten, welche den Verlauf dokumentieren und welche die abschließende ärztliche Beurteilung enthalten. Die Anwendung muss diese unterschiedlichen Aufgaben auch dann erkennen lassen, wenn sie dieselben Quellen verwendet.

Gemeinsame Informationen nutzen und Anforderungen der Kostenträger erhalten

Eine Rehaklinik behandelt Patient:innen zulasten der Rentenversicherung, der gesetzlichen Krankenversicherung und privater Versicherungen. Vorbefunde und ärztliche Festlegungen werden in allen drei Bereichen benötigt; Anforderungen an Berichte, Nachweise und Abrechnung können sich unterscheiden. Bei der Softwareauswahl ist deshalb zu klären, ob dieselben Angaben mehrfach gepflegt werden müssen und wie die jeweils geltenden Anforderungen zugeordnet werden. Auch eine spezialisierte Anwendung kann gut passen, wenn sie Informationen zuverlässig an die übrigen Systeme übergibt.

Ein umfangreiches Angebot aus einer Hand kann dieselben Abstimmungsprobleme enthalten: Wenn das Aufnahmemodul eine Korrektur speichert, die Dokumentation sie aber nicht übernimmt, entstehen auch innerhalb eines Produkts widersprüchliche Fassungen. Ein durchgespielter Änderungsfall liefert hier mehr Aufschluss als eine Funktionsübersicht. Die Klinik kann daran prüfen, ob Fallzuordnung, Änderungshistorie, Datenexport und Übergabe zwischen den Anwendungen zusammenpassen.

Eine US-Beobachtungsstudie untersuchte, ob die organisatorische Verbindung zwischen Akut- und Rehaklinik mit Unterschieden in der Aufenthaltsdauer einherging. Sood und Kolleg:innen verglichen innerhalb derselben Rehaeinrichtung die Aufenthaltsdauer von Patient:innen aus dem zugehörigen Akutkrankenhaus mit derjenigen von Patient:innen aus anderen Krankenhäusern. Bei räumlich angegliederter Rehabilitation waren die Aufenthalte von Patient:innen aus dem zugehörigen Akuthaus nach statistischer Anpassung kürzer. Bei räumlich

getrennten Einrichtungen desselben Verbunds zeigte sich kein vergleichbares Muster. Untersucht wurde die organisatorische Verbindung; ein Nutzen gemeinsamer Software lässt sich daraus nicht ableiten. Auch verbleibende Unterschiede zwischen den Patientengruppen können das Ergebnis mit erklären.¹³

Für die eigene Klinik lässt sich daraus eine konkrete Frage entwickeln: Bei welchen zuweisenden Häusern gelingt die Klärung offener Angaben besonders gut, und worauf beruht das? Vielleicht benennt eine Ärztin Unsicherheiten zuverlässig oder ein bestimmter Kollege ist bei Rückfragen erreichbar. Die Einführung einer gemeinsamen Anwendung bietet die Gelegenheit, solche Abläufe auch für Vertretungen zugänglich zu machen. Wiederholte Rückfragen nach zuweisender Stelle auszuwerten kann zeigen, ob Unterlagen fehlen, Zusagen unklar sind oder tatsächliche Versorgungsgrenzen erreicht werden.

Die für den einzelnen Aufenthalt gültige Regel muss erkennbar sein

Ein Entlassungsbericht kann im Dezember vorbereitet werden, obwohl die Entlassung erst im Januar erfolgt. Für die Klassifikation therapeutischer Leistungen (KTL) ist diese Unterscheidung bei der nächsten Umstellung relevant: Nach DRV-Vorgabe gilt die KTL 2025 für Entlassungen ab dem 1. Januar 2027. Eine automatische Prüfung muss daher die Fassung nach dem Entlassungsdatum zuordnen können. Die neueste Regeldatei im System zu hinterlegen genügt dafür als Verfahrensbeschreibung noch nicht.¹⁴

Bei örtlichen Vereinbarungen mit Kostenträgern stellt sich dieselbe Frage: Welcher Vertrag gilt für diesen Aufenthalt und welche Angaben verlangt er? KI kann Vertragsstellen und passende Falldaten leichter auffindbar machen. Ob die Voraussetzungen eines Anspruchs tatsächlich erfüllt sind, bleibt gesondert zu beurteilen. Eine gemeinsame Informationsgrundlage wird gerade dann nützlich, wenn sie solche Unterschiede zwischen Versorgungswegen erhält.

7. Wie KI-Korrekturen Lernen fördern und wiederkehrende Probleme verdecken können

Im folgenden Beispiel fehlen bei Anmeldungen aus einem Akutkrankenhaus wiederholt Operationsberichte. Eine KI-Prüfung kennzeichnet die Lücke; das

Aufnahmeteam fordert den Bericht an und erhält ihn noch rechtzeitig. Die einzelne Aufnahme kann dadurch besser vorbereitet werden. Für die Leitung bleibt jedoch eine zweite Frage: Wird inzwischen zuverlässig vollständig überwiesen, oder fängt das Team denselben Fehler nur schneller auf?

Wiederholte Nacharbeit als Anlass für eine Prozessänderung nutzen

Tucker und Edmondson beobachteten 26 Pflegefachpersonen während insgesamt 239 Stunden in neun Krankenhäusern. Sie untersuchten, wie die Beschäftigten mit Störungen ihrer Arbeit umgingen. Bei 93 Prozent der erfassten Probleme beschafften sie das unmittelbar Benötigte, ohne die Ursache weiter zu untersuchen oder weiterzugeben. Die ältere qualitative Studie liefert keine aktuelle Häufigkeit für Rehakliniken. Sie beschreibt aber, wie erfolgreiches Improvisieren dazu führen kann, dass ein wiederkehrender Mangel organisatorisch unbearbeitet bleibt.¹⁵

Für KI-gestützte Dokumentenprüfungen heißt das: Die Klinik braucht neben den Hinweisen zum einzelnen Fall eine Auswertung bestätigter, wiederkehrender Mängel. Im Beispiel wäre interessant, wie oft der Operationsbericht tatsächlich fehlte und ob dies nach einer Vereinbarung mit dem zuweisenden Haus seltener vorkommt. Die Zahl der KI-Warnungen allein wäre irreführend: Sie kann auch steigen, weil mehr Anmeldungen bearbeitet werden oder eine neue Modellfassung empfindlicher reagiert.

Die laufende Nacharbeit wird damit selbst zu einer Informationsquelle für das Qualitätsmanagement. Wiederholte Nachforderungen können eine Änderung der Übermittlungspraxis begründen; regelmäßig missverstandene Textstellen eine Überarbeitung der Vorlage. Ob die Änderung hilft, lässt sich daran prüfen, ob derselbe bestätigte Mangel bei den folgenden vergleichbaren Anmeldungen seltener auftritt. Wenn die Anwendung nur korrigierte Endergebnisse zeigt, bleibt der Aufwand für diese wiederholten Reparaturen leicht unsichtbar.

Ärztliche Urteilskraft lässt sich am fertigen Text schwerer beurteilen

Eine Assistenzärztin legt einen klar gegliederten Entlassungsentwurf vor, den sie mit KI vorbereitet hat. Die Oberärztin kann dessen Aussagen prüfen. Aus der sprachlichen Qualität lässt sich jedoch kaum erkennen, welche Zusammenhänge die Kollegin selbst erarbeitet und welche sie aus dem Modellvorschlag übernommen hat. Für die Weiter-

bildung verändert sich damit eine bisher beiläufige Informationsquelle: Die Bearbeitung des Textes macht Verständnislücken möglicherweise seltener sichtbar.

Zwei Studien von Goh und Kolleg:innen zeigen, weshalb der Zugang zu KI und die Qualität ärztlicher Bearbeitung getrennt betrachtet werden müssen. In der 2024 veröffentlichten Untersuchung lösten 50 Ärzt:innen diagnostische Fallvignetten mit GPT-4 und üblichen Hilfsmitteln oder nur mit den üblichen Hilfsmitteln. Bewertet wurden Diagnosen, Begründungen und vorgeschlagene Untersuchungen; ein statistisch gesicherter Vorteil durch KI zeigte sich nicht.¹⁶ In der 2025 veröffentlichten Untersuchung bearbeiteten 92 Ärzt:innen fünf simulierte Fälle zur weiteren Versorgung, ebenfalls mit GPT-4 und üblichen Hilfsmitteln oder nur mit den üblichen Hilfsmitteln. Die Bewertung von Versorgungsentscheidungen, diagnostischen Überlegungen und Fachwissen fiel mit KI bereinigt um 6,5 Prozentpunkte besser aus; die Bearbeitung dauerte rund zwei Minuten länger je Fall. Beide Studien prüften schriftliche Fallbearbeitung, keine Behandlungsergebnisse oder längerfristige Kompetenzentwicklung.¹⁷

Für ein Weiterbildungskonzept ist deshalb relevant, welche eigene Überlegung vor der KI-Rückmeldung festgehalten wird. Eine Ärztin kann etwa zunächst begründen, welche Befunde ihre abschließende Beurteilung tragen. Anschließend lässt sie mögliche Lücken oder Gegenargumente benennen und bespricht strittige Punkte mit einer erfahrenen Person. Das schafft einen Anlass, die eigene Begründung zu überprüfen. Ob diese Form der Zusammenarbeit die Urteilstkraft dauerhaft verbessert, ist damit noch offen.

Eine kleine randomisierte Untersuchung in Münster erprobte eine verwandte Lehrform mit Medizinstudierenden: Beide Gruppen führten vier simulierte Anamnesegespräche; eine erhielt zusätzlich Rückmeldungen eines Sprachmodells. Unter den 21 ausgewerteten Teilnehmenden verbesserte sich diese Gruppe stärker in der Bewertung durch verblindete Prüfende. Die Untersuchung zeigt einen kurzfristigen Nutzen in einer Lehrsituation. Langfristigen Kompetenzerhalt oder bessere selbstständige ärztliche Arbeit belegt sie nicht.¹⁸ Arbeitsproben ohne KI bleiben deshalb aufschlussreich, wenn eine Klinik beurteilen will, welche Fähigkeiten ihre Ärzt:innen auch ohne Unterstützung beherrschen.

Frei werdende Zeit gezielt für verständliche Patienteninformation einsetzen

Wenn KI bei der Zusammenstellung des Entlassungsberichts Arbeit abnimmt, kann die Klinik einen Teil der gewonnenen Zeit in ein zusätzliches Dokument investieren: einen ärztlich geprüften Patientenbrief. Er erläutert den Befund und die vereinbarten nächsten Schritte in verständlicher Sprache und greift die besprochenen Fragen zum Alltag auf. Damit wird aus dem Zeitgewinn eine zusätzliche Leistung. Ihr Erfolg wäre daran zu beurteilen, was Patient:innen mit der Erklärung anfangen können.

Rust und Kolleg:innen untersuchten dies bei 128 internistischen Patient:innen in Deutschland. Die zufällig zugeteilten Gruppen lasen entweder den üblichen Arztbrief oder ausgewählte, mit KI verständlich umformulierte und ärztlich geprüfte Abschnitte. Nach dem Lesen und vor dem Entlassgespräch bewerteten die Patient:innen ihr Wissen, ihre Zuversicht und ihre Fähigkeit zum Umgang mit der eigenen Gesundheit. Der Unterschied im Fragebogen PAM-13 betrug zugunsten der verständlichen Fassung 9,6 Punkte auf einer Skala von 0 bis 100. Objektives Wissen und langfristige Therapietreue wurden damit nicht nachgewiesen.¹⁹



Ein zusätzlicher Patientenbrief kann Befund und vereinbarte Schritte verständlich erläutern. Welche Erklärung im Alltag hilft, wird mit der betreffenden Person geklärt.

Ein Patientenbrief kann dokumentierte Wünsche, sprachliche Bedürfnisse und Lebensumstände aufgreifen. Aus Beruf oder Bildungsabschluss eine bevorzugte Erklärungsweise abzuleiten, bleibt jedoch unsicher. Eine Rückfrage zur gewünschten Darstellung schafft eine bessere Grundlage. Ob der Brief hilft, lässt sich anschließend daran prüfen, wie verständlich die Person ihre Situation und die vereinbarten Schritte erklären kann.

Klinische Unterlagen aufbereiten und Aufnahmen gemeinsam steuern

Mit aiomics lassen sich Vorbefunde und eigene Notizen für medizinische Dokumente nutzen. Im Aufnahmemanagement zeigt eine gemeinsame Übersicht, welche Anfragen vorliegen und wie weit ihre Bearbeitung ist. Wir passen die Vorlagen und Abläufe mit dem Team der jeweiligen Klinik an.

Unterlagen für mehrere Dokumente nutzen

Vorbefunde und eigene Aufzeichnungen können für Aufnahmeberichte, Entlassungsberichte und weitere fallbezogene Schreiben aufbereitet werden. Quellenbelege erleichtern den Abgleich mit den Unterlagen. Automatische Prüfungen unterstützen die Suche nach Widersprüchen; die fachliche Kontrolle bleibt erforderlich.

Anforderungen der Einrichtung abbilden

Unsere Bibliothek umfasst mehr als 100 Vorlagen. Wir passen sie an Fachgebiete, Empfänger:innen, Kostenträger und örtliche Vorgaben an. Die Beispiele in diesem Whitepaper sind bewusst verkürzt. Bei der Entwicklung von Prompts nutzen wir auch KI; die fachliche Erprobung bleibt Bestandteil der Arbeit.

Informationssicherheit verankern

Unser Informationssicherheitsmanagementsystem ist nach ISO/IEC 27001:2022 durch TÜV NORD CERT zertifiziert. Die vereinbarte Verarbeitung erfolgt in der EU; Patientendaten werden nicht zum Modelltraining verwendet. Der Einsatz wird in den Datenschutz- und Sicherheitsrahmen der Klinik eingebunden.

Den Nutzen mit der Klinik beurteilen

Die methodischen Grundlagen unseres Bewertungspakets wurden mit dem Charité-Team um Professor Akira Sebastian Poncette erarbeitet. Wir haben die Arbeitshilfen für unterschiedlich aufwendige örtliche Erhebungen weiterentwickelt. Das ermöglicht eine nachvollziehbare Bewertung anhand eigener Daten und ausdrücklich benannter Schätzungen.

Gern zeigen wir an einem konkreten Reha-Ablauf, welche Unterstützung bereits möglich ist und welche Anpassungen Ihre Einrichtung benötigt.

Unverbindlich zeigen lassen

8. Wann weniger Dokumentationsarbeit einen wirtschaftlichen Vorteil bringt

Ein Anbieter zeigt, wie aus einer Patientenakte in kurzer Zeit ein Entlassungsentwurf entsteht. Die Klinik muss anschließend Quellen prüfen, Aussagen korrigieren und den Bericht freigeben. Für die Investitionsentscheidung zählt daher, wie viel Arbeit bis zum verwendbaren Ergebnis anfällt und wo sie eingespart wird. Erst dann lässt sich klären, ob daraus weniger Mehrarbeit, zusätzliche Versorgung oder eine tatsächlich geringere Ausgabe entsteht.

Den Aufwand bis zum freigegebenen Bericht vergleichen

Nakano und Imura untersuchten in einem japanischen Krankenhaus therapeutische Übergabezusammenfassungen für die Rehabilitation. Verglichen wurden 16 Berichte vor und 22 nach Einführung einer KI-gestützten Erstellung. Die mediane Arbeitszeit einschließlich Prüfung, Korrektur und Fertigstellung sank von 23 auf 10 Minuten. In einer gesonderten Stichprobe enthielten neun von elf KI-Entwürfen mindestens einen Fehler. Der Zeitvorteil trat also trotz notwendiger Nacharbeit auf. Die kleine Untersuchung ohne gleichzeitige Vergleichsgruppe betraf Therapeut:innen und erlaubt keine allgemeine Prognose für ärztliche Reha-Berichte in Deutschland.²⁰

Tan und Kolleg:innen beobachteten 169 ambulante Konsultationen bei neun erfahrenen Ärzt:innen in Singapur, mit und ohne gesprächsbegleitende Dokumentationshilfe. Sie erfassten auch Prüfung, Bearbeitung und manuelle Übertragung der Texte. Die statistisch angepasste Dokumentationszeit sank mit der Anwendung um 0,8 Minuten; der Anteil der Konsultationszeit mit Blickkontakt stieg um 7,4 Prozentpunkte. Die Gesamtdauer veränderte sich nicht statistisch signifikant. Die kleine, nicht randomisierte Untersuchung belegt keine Personalsparnis im stationären Bereich. Sie zeigt, dass weniger Dokumentationsarbeit auch der Zuwendung im bestehenden Termin zugutekommen kann.²¹

Festlegen, wofür eingesparte Arbeitszeit verfügbar wird

Werden Entlassungsberichte innerhalb der regulären Dienstzeit fertig, kann dies Mehrarbeit verringern. Entfällt eine sonst erforderliche zusätzliche Besetzung, kann eine Ausgabe vermieden werden. Werden mehr Fälle aufgenommen, entsteht ein zusätzlicher Deckungsbeitrag nur, soweit die Klinik

auch Pflege, Therapie und weitere benötigte Leistungen bereitstellen kann. Diese Verwendungen eingesparter Zeit haben unterschiedliche wirtschaftliche Voraussetzungen. Für die Rechnung muss erkennbar sein, welche davon die Klinik tatsächlich erreichen will.

In einem Haus mit Akut- und Rehabilitationsversorgung kann dieselbe organisatorische Verbesserung wirtschaftlich verschieden wirken. Werden durch bessere Abstimmung Entlassungen früher möglich, sinkt bei tagessatzvergüteter Rehabilitation zunächst auch die Zahl abrechenbarer Tage. Bei fallbezogener Vergütung können unter den geltenden Bedingungen Kosten sinken. Ob und zu welchen Bedingungen die frei gewordene Kapazität zusätzliche Leistung ermöglicht, gehört jeweils gesondert in die Rechnung. Eine einheitliche Wirtschaftlichkeitsannahme für das gesamte Haus kann diese Unterschiede verdecken.

Auch der ab Januar 2027 veränderte Prüfquotenrahmen nach § 275c SGB V ist nach Leistungsbe- reich einzuordnen. Er betrifft Krankenhausbehandlung und damit entsprechende Akutleistungen eines gemischten Hauses. Für DRV-Rehabilitation entsteht daraus keine allgemeine Prüfquote. Wenn eine Anwendung Nachweise aus der Patientenakte zusammenführt, kann sie beide Bereiche unterstützen; der wirtschaftliche Nutzen richtet sich nach den jeweils geltenden Nachweis- und Vergütungsbedingungen.²²

Ein weiterer möglicher Nutzen betrifft gesondert erstattungsfähige Arzneimittelkosten. Eine Anwendung kann beispielsweise dokumentierte Medikamentengaben mit den für den Aufenthalt geltenden Vereinbarungen abgleichen und Positionen zur Prüfung vorlegen. In die Nutzenrechnung gehen die zusätzlich realisierbaren Beträge nach Abzug des Bearbeitungsaufwands ein. Bereits regelmäßig geltend gemachte Positionen bringen keinen zusätzlichen Erlös. Ob eine gesonderte Erstattung möglich ist, hängt vom Kostenträger und vom konkreten Vertrag ab.^{23,24}

Auf der Kostenseite werden die Arbeitsschritte derselben Fälle erfasst: Unterlagen übertragen, Entwürfe gegenlesen, Fehler korrigieren und die Anwendung betreuen. Enthält die gemessene Zeiterparnis diese Arbeit bereits, würde ein erneuter Abzug die Bilanz verzerren. Umgekehrt wäre es eine Doppelzählung, dieselbe freie Stunde vollständig als vermiedene Personalausgabe und zugleich als Arbeitszeit für zusätzliche Behandlungen anzusetzen.

Die Annahme prüfen, an der die Investitionsentscheidung hängt

Das folgende Rechenbeispiel ist frei gewählt. Eine Klinik erstellt 200 Entlassungsberichte im Monat. Unklar ist, ob die KI nach Einbeziehung aller Prüf- und Korrekturschritte insgesamt sechs, neun oder zwölf Minuten je Bericht einspart. Eine Arbeits-

stunde wird intern mit 80 Euro bewertet. Für die Rechnung wird außerdem angenommen, dass die Hälfte des Zeitgewinns eingeplante zusätzliche Arbeitsstunden und damit eine benannte Ausgabe vermieden. Dieser Anteil braucht eine örtliche Begründung; ohne tatsächlich vermeidbare Ausgabe kann er null betragen.

TABELLE 1

Beispiel: Wann die monatlichen Kosten gedeckt wären

Annahme und Ergebnis	Vorsichtig	Mittlerer Ansatz
Eingesparte Gesamtarbeitszeit je Entlassungsbericht	6 Minuten	9 Minuten
Wert freier Arbeitszeit bei 200 Berichten und 80 Euro je Stunde	1.600 Euro	2.400 Euro
Tatsächlich vermiedene Ausgabe bei angenommenen 50 Prozent	800 Euro	1.200 Euro
Gesamtkosten je Monat einschließlich anteiliger Einführung	1.200 Euro	1.200 Euro
Monatlicher Saldo	-400 Euro	0 Euro

Bei zwölf eingesparten Minuten je Bericht läge der monatliche Saldo unter denselben Annahmen bei plus 400 Euro. Bei neun Minuten wären die monatlichen Kosten gedeckt. Der nächste Erhebungsschritt kann sich damit auf die Größe konzentrieren, die den Beschluss noch verändert: etwa die umstrittene Prüf- und Korrekturzeit oder die Frage, ob zusätzliche Arbeitsstunden tatsächlich entfallen. Eine allgemeine Zufriedenheitsbefragung würde diese konkrete Unsicherheit kaum klären.

Für die Hochrechnung werden die eingesparten Personenminuten aller betrachteten Berichte summiert oder ein geeigneter arithmetischer Mittelwert verwendet. Wenige sehr aufwendige Berichte können die durchschnittliche Arbeitszeit stark erhöhen, während sich der Median kaum verändert. Eine Hochrechnung allein mit dem Median würde diesen Aufwand dann unterschätzen. Außerdem zählt, bei wie vielen der geeigneten Berichte die KI tatsächlich genutzt wird. Der Vorteil aus ausgewählten, gut dokumentierten Fällen lässt sich ohne Prüfung nicht auf das gesamte monatliche Arbeitsaufkommen übertragen.

Die begleitende Rechenvorlage hält den rechnerischen Wert freier Arbeitszeit, tatsächlich vermiedene Ausgaben und Kosten getrennt. Sie berücksichtigt auch, wenn Arbeit zwischen Berufsgruppen verlagert wird; zusätzliche Deckungsbeiträge sind gesondert zu bewerten. Die enthaltenen Beispielwerte sind veränderbar. Eingesetzte Schätzungen bleiben als solche zu kennzeichnen.

9. Wie Kliniken den Nutzen vor einer Ausweitung beurteilen können

Nach einer Erprobung der KI-gestützten Entlassungsdokumentation steht eine Klinik vor der nächsten Entscheidung. Soll die Anwendung in einer Abteilung weiterlaufen, auf mehrere Häuser ausgeweitet oder mit einer öffentlich bezifferten Einsparung beworben werden? Dafür werden unterschiedlich belastbare Daten benötigt. Der Erhebungsaufwand sollte davon abhängen, welche dieser Entscheidungen ansteht und welche Folgen eine falsche Einschätzung hätte.

Dafür werden im Folgenden zwei Wege beschrieben. Der erste plant einen Vergleich mit gemesse-

nen Zeiten und geprüfter Dokumentenqualität, bevor die Erhebung beginnt. Der zweite verbindet vorhandene Daten mit ausdrücklich ausgewiesenen Schätzungen des Teams, um einen begrenzten nächsten Schritt zu beurteilen. Beide verwenden Bausteine aus dem Einführungs- und Bewertungspaket von aiomics. Die vereinfachte Erhebung lässt die fachliche Prüfung und Freigabe der erzeugten Dokumente unberührt.

Weg A: Den Nutzen mit einem geplanten Vergleich untersuchen

Wenn Entlassungsberichte nach der KI-Einführung früher fertig werden, können mehrere Ursachen beteiligt sein: die Anwendung, eine veränderte Besetzung oder weniger komplexe Fälle. Ein geeigneter Vergleich hilft, diese Einflüsse auseinanderzuhalten. Infrage kommen eine zufällige Zuteilung vergleichbarer Bereiche oder Einführungszeitpunkte, eine begründete gleichzeitig beobachtete Vergleichsgruppe oder eine ausreichend lange Zeitreihe. Die genaue Messung vor und nach der Einführung allein trennt die Wirkung der KI noch nicht von anderen Veränderungen.

Bei einer schrittweisen Einführung in mehreren Abteilungen lässt sich die Reihenfolge unter geeigneten Bedingungen zufällig zuteilen. Die Auswertung muss dann sowohl Unterschiede zwischen den Abteilungen als auch Entwicklungen über die Zeit berücksichtigen. Eine bloß zeitlich gestaffelte Einführung erfüllt diese Anforderungen an eine randomisierte Untersuchung noch nicht. Wie viele Fälle und welcher Zeitraum benötigt werden, hängt unter anderem von Zielgröße, Streuung und erwarteter Wirkung ab; die methodische Planung muss diese Größen aufeinander abstimmen.²⁵

Vor der Erhebung wird festgehalten, welche Entlassungsberichte für die Unterstützung geeignet sind und womit die KI-gestützte Bearbeitung verglichen wird. Auch geeignete Berichte, bei denen niemand die Anwendung nutzt, eine Nutzung abbricht oder die Freigabe noch aussteht, werden erfasst. So bleiben zwei Ergebnisse unterscheidbar: der Nutzen bei tatsächlicher Anwendung und der Nutzen über das gesamte dafür vorgesehene Arbeitsaufkommen. Für eine Ausweitung ist gerade der zweite Wert relevant.

Zur Arbeitszeit gehören die Beiträge aller beteiligten Berufsgruppen bis zur Freigabe. Ein Zeitstempel zeigt etwa, wann ein Entwurf angelegt und wann er unterschrieben wurde; dazwischen liegen auch Wartezeiten. Aktive Arbeitsminuten müssen des-

halb gesondert erfasst oder aus dafür geeigneten Daten bestimmt werden. Für die Dokumentenqualität werden vorab Kriterien festgelegt. Prüfende, die den jeweiligen Bericht nicht selbst bearbeitet haben, beurteilen ihn anhand der Quellen; bei abweichenden Urteilen braucht es ein festgelegtes Klärungsverfahren.

Entwurf, Korrekturarbeit und freigegebener Bericht werden getrennt betrachtet. So kann sichtbar werden, dass die Ärzt:innen erhebliche Fehler zuverlässig auffangen, dafür aber viel Zeit benötigen. Wenige dokumentierte Korrekturen können umgekehrt sowohl auf gute Entwürfe als auch auf übersehene Mängel zurückgehen. Eine Fehlerquote wird deshalb zusammen mit der Zahl und Auswahl der geprüften Berichte sowie dem verwendeten Prüfverfahren berichtet.

Zu einer wissenschaftlichen Untersuchung gehören außerdem der Auswertungsplan, der Umgang mit fehlenden Angaben und die erforderlichen ethischen und datenschutzrechtlichen Klärungen. Modell, Prompt und technische Verarbeitung werden mit ihren eingesetzten Fassungen dokumentiert. Für die Berichterstattung früher klinischer KI-Untersuchungen bietet DECIDE-AI eine Struktur; für Untersuchungen zur Qualitätsverbesserung kommt SQUIRE 2.0 infrage. Beide helfen offenzulegen, was tatsächlich untersucht wurde. Die Eignung des gewählten Vergleichs bleibt unabhängig davon zu begründen.^{26,27}

Weg B: Eine begrenzte Entscheidung mit ausgewiesenen Schätzungen vorbereiten

Eine Abteilung will zunächst entscheiden, ob sie die Dokumentationshilfe für einen begrenzten Zeitraum weiterverwendet. Für eine umfassende Zeiterhebung fehlen die Ressourcen. Dann können vorhandene Berichtszahlen, kurze Einschätzungen anhand kürzlich bearbeiteter Fälle und gezielte Gegenprüfungen eine Entscheidungsgrundlage liefern. Vorab werden der betrachtete Bereich, die Art der Berichte und der nächste Beschluss benannt. Erinnernte Bearbeitungszeiten gehen ausdrücklich als Schätzungen ein.

Die beteiligten Berufsgruppen schätzen ihren Aufwand zunächst getrennt. Ärzt:innen können etwa weniger Zeit für den ersten Entwurf, Oberärzt:innen zugleich mehr Zeit für die Freigabe angeben. Die Schätzungen beziehen sich auf kürzlich bearbeitete Berichte und enthalten einen plausiblen unteren und oberen Wert. Erst danach werden die Angaben gemeinsam besprochen. So bleibt erkennbar, ob

die Beteiligten tatsächlich unterschiedliche Arbeitsschritte sehen, statt sich vorschnell auf einen Mittelwert zu einigen.

Vorhandene Daten begrenzen den Spielraum der Schätzung: Zahl geeigneter und tatsächlich unterstützter Berichte, Rechnungen, Freigabezeitpunkte und dokumentierte Rückgaben zur Überarbeitung. Entscheidet die oberärztliche Prüfzeit über die Wirtschaftlichkeit, kann sie an einigen Fällen gezielt beobachtet werden. Schon wenige Beobachtungen können einen bislang vergessenen Arbeitsschritt sichtbar machen. Für eine genaue Schätzung seiner durchschnittlichen Dauer reichen sie möglicherweise noch nicht.

In der Rechnung wird jeder Wert als gemessen, geschätzt oder angenommen gekennzeichnet. Eine Spanne aus Einschätzungen zeigt, wie unsicher sich die Beteiligten sind; sie ist kein statistisches Konfidenzintervall. Fehlt eine Messung aus der Zeit vor der Einführung, wird diese Lücke benannt. Frühere Aufzeichnungen können einen gemessenen Ausgangswert liefern. Die rückblickende Erinnerung des Teams bleibt eine geschätzte Ausgangslage.

Auch bei diesem Weg werden Berichte fachlich freigegeben und kritische Vorfälle bearbeitet. Die Auswertung kann gemeldete Probleme berücksichtigen; unbemerkte Fehler erfasst sie damit nicht zuverlässig. Weder eine günstige Zeitrechnung noch einige fehlerfreie Berichte belegen klinische Sicherheit oder Gleichwertigkeit.

TABELLE 2

Welche Schlussfolgerung die beiden Bewertungswege erlauben

Anstehende Entscheidung	Geplanter Vergleich mit Messung	Pragmatische Bewertung
Begrenzte Fortführung	Nutzen und Unsicherheit nach Vergleichsverfahren und Datenumlage beurteilen.	Möglich, wenn vorsichtige Annahmen tragen und wesentliche Qualitätsbedenken geklärt sind.
Einführung in weiteren Bereichen	Übertragbarkeit auf deren Fälle, Personal und Abläufe untersuchen.	Kann nächste Prüfschritte begründen; örtliche Schätzungen sichern die Übertragung noch nicht ab.
Öffentlich bezifferte Wirkung	Nur so weit beziffern, wie Vergleich, Messung und Auswertung die Aussage tragen.	Als örtliche Einschätzung mit Annahmen darstellen; keine durch die Anwendung verursachte Einsparung behaupten.
Klinische Sicherheit oder Gleichwertigkeit	Dafür geeignete Kriterien, Fallzahlen und Auswertung eigens vorsehen.	Aus Zeit- und Nutzenschätzungen nicht ableitbar.

Die nächste Messung auf die offene Entscheidungsfrage richten

Im Rechenbeispiel des vorigen Kapitels liegt die Schwelle bei neun eingesparten Minuten je Entlassungsbericht. Umfassen die plausiblen Annahmen Werte darunter und darüber, kann eine gezielte Zeitmessung den Beschluss verändern. Trägt dagegen bereits die vorsichtige Rechnung eine kleine, rücknehmbare Fortführung, kann eine begrenzte Erhebung angemessen sein. Eine größere Investition oder eine weitreichende Veränderung klini-

scher Verantwortung rechtfertigt einen höheren Aufwand für die Absicherung.

Geht es um die Frage, weshalb eine Anwendung wenig genutzt wird, kann eine Befragung geeigneter sein als eine weitere Zeitmessung. Die Fragebögen AIM, IAM und FIM erfassen Akzeptanz, wahrgenommene Angemessenheit und Durchführbarkeit; ihre deutschen Fassungen wurden psychometrisch untersucht. Wer daraus einzelne Fragen zu einer eigenen Kurzbefragung zusammenstellt, verändert jedoch das Instrument. Die Prüfung der

ursprünglichen Skalen validiert diese neue Kombination nicht automatisch.²⁸

Die beiden Bewertungsblätter am Ende helfen, Entscheidung, Datenquellen, Annahmen und offene Fragen festzuhalten. Für Weg A werden zusätzlich ein vollständiges Untersuchungsprotokoll und eine geeignete Auswertung benötigt. Bei Weg B dokumentieren Schätztabelle, verfügbare Mengenangaben, bekannte Qualitätsprobleme und begründeter Beschluss die Entscheidungsgrundlage. Nach außen wird erkennbar berichtet, ob die Klinik Zeiten gemessen oder den Nutzen anhand örtlicher Schätzungen beurteilt hat.

10. Wie ein begrenzter Einstieg die spätere Einführung vorbereitet

Eine Klinik kann die KI-gestützte Entlassungsdokumentation zunächst in einem Bereich erproben, bevor sie weitere Abteilungen anschließt. Dabei wird sichtbar, welche Unterlagen sich gut verarbeiten lassen, wie viel Prüfung die Ärzt:innen benötigen und wo zusätzliche Übertragungsarbeit anfällt. Für die spätere Einführung sind gerade die schwierigen Fälle aufschlussreich. Eine Erprobung mit ausschließlich vollständigen, gut lesbaren Akten würde ihren Aufwand kaum erkennen lassen.

Beim begrenzten Einstieg auch die spätere Übernahme der Daten prüfen

Bei einer Anwendung im Browser können Beschäftigte Unterlagen zunächst gezielt hochladen und freigegebene Texte in das Krankenhausinformationssystem übernehmen. So lässt sich der Nutzen unter Umständen vor einer vollständigen technischen Anbindung beurteilen. Gleichzeitig entsteht manuelle Arbeit für Übertragung und Zuordnung. In die Entscheidung über diesen Einstieg gehören deshalb sowohl der kurzfristige Nutzen als auch die Möglichkeiten einer späteren Anbindung.

Während der Erprobung entstehen außerdem angepasste Prompts, Bearbeitungsvermerke und Korrekturen zu einzelnen Fällen. Die Klinik sollte an Beispieldaten prüfen können, welche dieser Informationen sie in nutzbarer Form exportieren und bei einem Anbieterwechsel weiterverwenden kann. Mit jeder zusätzlichen Vorlage und jeder eingebundenen Abteilung wächst der mögliche Aufwand für einen Wechsel. Diese künftigen Kosten können relevant sein, obwohl sie in der ersten Monatsrechnung noch fehlen.

Die beobachtete Übertragungsarbeit hilft auch, Schnittstellen zu priorisieren. Muss das Aufnahmeteam täglich Faxe manuell zuordnen, kann dort ein früherer Nutzen liegen als bei einer selten verwendeten Komfortfunktion. Werden externe Befunde wiederholt abgetippt, lohnt sich die Prüfung einer direkten Übernahme. Zu jedem Anschluss gehören neben technischer Machbarkeit und Vertrag auch erkennbare Fehlerfälle: Wer bemerkt beispielsweise eine fehlgeschlagene Übertragung und wie wird sie nachgeholt?

Neue Funktionen verlangen eine erneute Prüfung ihres Einsatzes

Eine Anwendung bereitet zunächst vorhandene Angaben für einen Entlassungsbericht auf. Später soll sie aus denselben Daten patientenspezifische Empfehlungen für Diagnostik oder Behandlung ableiten. Damit verändert sich auch die Frage der medizinprodukterechtlichen Einordnung. Maßgeblich sind Zweckbestimmung und tatsächliche Funktion. Für Software, die als Medizinprodukt einzuordnen ist, regelt Regel 11 der Medizinprodukteverordnung die risikobezogene Klassifizierung. Eine abschließende ärztliche Freigabe begründet für sich genommen keine Ausnahme.²⁹

Ändern sich Funktionen oder Nutzerkreis, sind auch Datenschutz und die Beteiligung der zuständigen Beschäftigtenvertretung auf den tatsächlichen Einsatz zu beziehen. Schon ein begrenzter Betrieb ohne Schnittstelle verarbeitet sensible Patientendaten; gespeicherte Bearbeitungsstände können zugleich Arbeitsabläufe nachvollziehbar machen. Die Orientierungshilfe der Datenschutzkonferenz beschreibt Anforderungen an Auswahl und Nutzung von KI-Anwendungen. Eine wesentlich veränderte Nutzung gibt Anlass, die dafür getroffenen Regelungen erneut zu prüfen.³⁰

Festhalten, welche Veränderung die nächste Bewertung auslöst

Hat die Erprobung einen Nutzen gezeigt, lässt sich die Fortführung an benannte Bedingungen knüpfen. Die geprüfte Modellfassung, die verwendeten Prompts und die einbezogenen Fallgruppen gehören dazu. Ein Modellwechsel, eine neue Fachabteilung oder eine veränderte Aufgabenverteilung kann dann gezielt eine erneute Bewertung auslösen. Die Klinik muss dafür wissen, auf welchen Bedingungen ihr bisheriger Beschluss beruhte.

BEWERTUNGSBLATT 1

Entscheidung und Datenlage

Für eine begrenzte Fortführung, eine Ausweitung, einen Vergleich oder eine öffentliche Wirkungsaussage. Jede Zahl behält ihre Herkunft und ihren Geltungsbereich.

Entscheidung und Reichweite

Beschluss bis / finanzieller Rahmen

Vorgehen: Messung / kurze Bewertung

Anwendung / beteiligte Bereiche

Vergleichszeiträume / maßgebliche Fassung

Geeignete Fälle / tatsächliche Nutzung

Ergebnisse mit erkennbarer Herkunft

Ergebnis	Herkunft ¹	Bisher	Danach / erwartet	Fälle / Zeitraum	Unsicherheit
Geeignete Fälle und tatsächliche Nutzung					
Aktive Arbeitszeit nach Berufsgruppe					
Zeit bis Entscheidung oder Abschluss					
Geprüfte Fehler und Prüfaufwand					
Erlöse / vermiedene Ausgaben					

¹ G = gemessen, S = geschätzt, A = angenommen. Zusätzlich Quelle und Bezugseinheit (Fall, Dokument, Woche, Befragte) angeben. Quoten als Zähler/Nenner; offene Fälle, Nichtnutzung und Fehlversuche einbeziehen.

Andere mögliche Erklärungen: Personal, Fallzusammensetzung, parallele Änderungen

Anlage bei Messung: Prüfplan und Auswertung. Bei kurzer Bewertung: Herkunft und Bandbreite der Schätzungen. Wesentliche Unterschiede zwischen Berufsgruppen gesondert festhalten.

BEWERTUNGSBLATT 2

Unsicherheit und Beschluss

Die Bandbreite zeigt, welche Annahme den Beschluss beeinflusst. Wirtschaftliche Vorteile und offene wesentliche Qualitätsfragen werden getrennt beurteilt.

Wirtschaftliche Bandbreite im gewählten Zeitraum

Größe	Vorsichtig	Günstig	Tragende Annahme
Verfügbare Arbeitskapazität (Stunden)			
Tatsächlich vermiedene Ausgaben (€)			
Zusätzlicher Deckungsbeitrag (€)			
Einführung und laufender Betrieb (€)			
Saldo zahlungswirksamer Effekte (€)			

Kapazität bleibt gesondert ausgewiesen. G = gemessen, S = geschätzt, A = angenommen. Keine doppelte Anrechnung derselben Zeit oder desselben wirtschaftlichen Effekts.

Annahme, die den Beschluss ändern könnte

Belastbare Klärung: Aufwand und Verzögerung

Offene wesentliche Qualitätsfrage

Beibehaltene Absicherung / nötige Begrenzung

Beschluss: begrenzt fortführen · weiter messen · anpassen · beenden

Reichweite des Beschlusses

Nächste Prüfung / Auslöser bei Änderungen

Durch die Daten gestützte Aussage

Von dieser Bewertung nicht gedeckte Aussage

Beschlussverantwortung und Datum: _____

Einzubeziehende Änderungen: Modell, Vorlagen, Quellenaufbereitung und maßgebliche Regeln. Die Bewertung gilt für den dokumentierten Einsatz und die festgehaltenen Annahmen.

Literatur und Quellen

Gezielte Auswertung ausgewählter Primärquellen; keine systematische Übersichtsarbeit und keine externe wissenschaftliche Begutachtung. Recherche- und Rechtsstand: 13. September 2026. Die Titel fremdsprachiger Veröffentlichungen wurden sinngemäß ins Deutsche übertragen; die Verweise führen zu den Originalquellen. Untersuchungsbe- funde, eigene Analysen, Beispiele und Unternehmensangaben sind im Text unterschieden. Die Bewertungsblätter sind eigenständige Anpassungen von Bausteinen des aiomics-Einführungspakets; sie ersetzen kein Studienprotokoll.

1. Neferu R, Fleck R (2024). Wartezeiten bis zur stationären Rehabilitation bei erworbener Hirnschädigung verkürzen: ein Projekt zur Qualitätsverbesserung. BMJ Open Quality 13:e002915. Titel ins Deutsche übertragen. DOI: 10.1136/bmjopen-2024-002915. [Quelle öffnen](#)
2. Sood N, Huckfeldt PJ, Grabowski DC, Newhouse JP, Escarce JJ. Auswirkungen prospektiver Vergütung auf Aufnahmeentscheidungen und Behandlungsintensität in stationären Rehabilitationseinrichtungen. Titel sinngemäß übersetzt. Journal of Health Economics. 2013;32(5):965–979. DOI: 10.1016/j.jhealeco.2013.05.003. [Quelle öffnen](#)
3. Deutsche Rentenversicherung. Fragen zur Einrichtungsauswahl: sozialmedizinische Kriterien, Qualität, Entfernung, Wartezeit sowie Wunsch- und Wahlrecht. Abruf am 13. September 2026. [Quelle öffnen](#)
4. Rolls D, Khanna S, Lloyd N, Reeson A, Jayasena R, McCormick C, Hakkennes S (2020). Verweildauer in der Rehabilitation vor und nach Einführung einer elektronischen Fallübersicht. International Journal of Medical Informatics 134:104042. Titel ins Deutsche übertragen. DOI: 10.1016/j.ijmedinf.2019.104042. [Quelle öffnen](#)
5. Mebrahtu TF, McInerney CD, Benn J und weitere (2023). Auswirkungen einer zentralen Krankenhaussteuerung auf Patientenfluss und Datenqualität: Ergebnisse aus dem britischen National Health Service. International Journal for Quality in Health Care 35:mzad072. Titel ins Deutsche übertragen. DOI: 10.1093/intqhc/mzad072. [Quelle öffnen](#)
6. Weigl M, Müller A, Vincent C, Angerer P, Sevdalis N (2012). Zusammenhang zwischen Arbeitsunterbrechungen und der Belastung von Krankenhausärzt:innen: eine prospektive Beobachtungsstudie. BMJ Quality & Safety 21:399–407. Titel ins Deutsche übertragen. DOI: 10.1136/bmjqs-2011-000188. [Quelle öffnen](#)
7. Williams CYK u. a. Ärztlich und mit Sprachmodellen erstellte stationäre Entlasszusammenfassungen. JAMA Internal Medicine. 2025;185(7):818–825. Verblindeter Vergleich von 100 stationären Verläufen. DOI: 10.1001/jamainternmed.2025.0821. [Quelle öffnen](#)
8. Omar M u. a. Anfälligkeit mehrerer Sprachmodelle für gezielt erfundene Angaben in der klinischen Entscheidungsunterstützung. Communications Medicine. 2025;5:330. Belastungsversuch mit 300 synthetischen Fällen. DOI: 10.1038/s43856-025-01021-3. [Quelle öffnen](#)
9. Wu D u. a. NOHARM: medizinische Sicherheitsbewertung und randomisierte Untersuchung der Zusammenarbeit von Ärzt:innen und KI. Vorabveröffentlichung arXiv:2512.01241, Fassung 4 vom 13. Juli 2026; noch nicht abschließend begutachtet. Konsultationsfälle aus zehn Fachgebieten; Bewertung möglicher Schäden. [Quelle öffnen](#)
10. Grolleau F u. a. Ärztliche Sicherheitsbewertungen KI-generierter Zusammenfassungen stationärer Verläufe. JAMA Network Open. 2026;9(5):e2616556. Prospektives Qualitätsverbesserungsprojekt mit 384 Entlassungen; ärztliche Rückmeldungen zu 100 Entwürfen. DOI: 10.1001/jamanetworkopen.2026.16556. [Quelle öffnen](#)
11. Deutsche Rentenversicherung. Der ärztliche Reha-Entlassungsbericht. Leitfaden, Stand März 2026; insbesondere S. 9–11, 27–34 und 37–46. Neue KTL-Regeln gelten für Entlassungen ab 1. Januar 2027. [Quelle öffnen](#)
12. Deutsche Rentenversicherung. Peer Review – Somatische Indikationen. Checkliste und Manual zur Prüfung der Prozessqualität anhand anonymisierter Entlassungsberichte und individueller Therapiepläne. Stand 21. Juni 2017. [Quelle öffnen](#)
13. Sood N, Shier V, Huckfeldt PJ, Weissblum L, Escarce JJ. Auswirkungen vertikal integrierter Versorgung auf Inanspruchnahme und Ergebnisse in stationären Rehabilitationseinrichtungen. Titel sinngemäß übersetzt. Health Services Research. 2021;56(5):828–838. DOI: 10.1111/1475-6773.13667. [Quelle öffnen](#)
14. Deutsche Rentenversicherung. Leitfaden zum einheitlichen Entlassungsbericht: Umstellung auf KTL 2025 zum 1. Januar 2027, maßgeblich nach Entlassungsdatum. Erläuterungen zur Fassung vom März 2026. [Quelle öffnen](#)
15. Tucker AL, Edmondson AC. Warum Krankenhäuser aus Störungen zu wenig lernen: organisatorische und psychologische Hindernisse für systematische Verbesserungen. Titel sinngemäß übersetzt. California Management Review. 2003;45(2):55–72. DOI: 10.2307/41166165. [Quelle öffnen](#)
16. Goh E u. a. Einfluss eines Sprachmodells auf das diagnostische Denken: randomisierte klinische Studie. JAMA Network Open. 2024;7(10):e2440969. DOI: 10.1001/jamanetworkopen.2024.40969. Untersuchung November/Dezember 2023. [Quelle öffnen](#)
17. Goh E u. a. KI-Unterstützung bei ärztlichen Versorgungsaufgaben: randomisierte Studie. Nature Medicine. 2025;31:1233–1238. DOI: 10.1038/s41591-024-03456-y. Untersuchung November 2023 bis April 2024. [Quelle öffnen](#)

18. Brügge E u. a. Klinisches Denken mit KI-simulierten Patientengesprächen und strukturierter Rückmeldung üben. BMC Medical Education. 2024;24:1391. Randomisierte Untersuchung in Münster; 21 Medizinstudierende ausgewertet. DOI: 10.1186/s12909-024-06399-7. [Quelle öffnen](#)
19. Rust P u. a. Wirkung patientenorientierter KI-Entlasszusammenfassungen auf die Patient:innenaktivierung: randomisierte Studie in Deutschland. The Lancet Digital Health. 2026;8(6):100991. DOI: 10.1016/j.landig.2026.100991. [Quelle öffnen](#)
20. Nakano M, Imura T. KI-gestützte Erstellung therapeutischer Entlasszusammenfassungen in der Rehabilitation: Zeitaufwand, Gebrauchstauglichkeit und Fehler. Cureus. 2026;18(5):e109247. Kleine Vorher-nachher-Untersuchung in Japan: 16 Berichte vor und 22 nach Einführung; gesonderte Fehlerprüfung an elf Entwürfen. DOI: 10.7759/cureus.109247. [Quelle öffnen](#)
21. Tan JYE, Rafi IBM, Sng GGR und weitere. Auswirkungen einer gesprächsbegleitenden KI-Dokumentationshilfe auf Ärzt:innen und Patient:innen: prospektive Beobachtungsstudie mit unmittelbarer Zeitmessung. Titel sinngemäß übersetzt. JMIR Medical Informatics. 2026;14:e85580. DOI: 10.2196/85580. [Quelle öffnen](#)
22. GKV-Beitragssatzstabilisierungsgesetz vom 24. Juli 2026, BGBl. 2026 I Nr. 228 vom 29. Juli 2026. Artikel 1 Nr. 66 und Artikel 8 Abs. 2: Änderung der Prüfquoten nach § 275c SGB V zum 1. Januar 2027; betrifft Krankenhausbehandlung. [Quelle öffnen](#)
23. Deutsche Rentenversicherung Schwaben. Hinweise für Rehabilitationseinrichtungen: Leistungen zur medizinischen Rehabilitation ohne Abhängigkeitserkrankungen. Version 9, S. 45–46: Medikamente; gesonderte Erstattung unter den dort genannten Voraussetzungen. Abruf am 13. September 2026. [Quelle öffnen](#)
24. GKV-Spitzenverband und Vereinbarungspartner. Rahmenempfehlungen Vorsorge und Rehabilitation nach §§ 111 Abs. 7 ff. SGB V vom 28. März 2025. Insbesondere Abschnitt 2.11 zum Entlassungsbericht und 3.2 Abs. 9 zu besonders teuren Arzneimitteln. [Quelle öffnen](#)
25. Hemming K, Taljaard M, McKenzie JE und weitere. Berichterstattung über cluster-randomisierte Studien mit gestaffelter Einführung: Erweiterung der CONSORT-Erklärung mit Erläuterungen. BMJ. 2018;363:k1614. Titel ins Deutsche übertragen. DOI: 10.1136/bmj.k1614. [Quelle öffnen](#)
26. Vasey B, Nagendran M, Campbell B und weitere. DECIDE-AI: Empfehlungen zur Berichterstattung über die frühe klinische Prüfung KI-gestützter Entscheidungshilfen. Nature Medicine. 2022;28:924–933. Titel ins Deutsche übertragen. DOI: 10.1038/s41591-022-01772-9. [Quelle öffnen](#)
27. Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. SQUIRE 2.0: Überarbeitete Empfehlungen zur Veröffentlichung von Untersuchungen zur Qualitätsverbesserung. BMJ Quality & Safety. 2016;25:986–992. Online veröffentlicht 2015. Titel ins Deutsche übertragen. DOI: 10.1136/bmjqs-2015-004411. [Quelle öffnen](#)
28. Kien C, Griebler U, Schultes MT, Thaler KJ, Stamm T. Psychometrische Prüfung der deutschen Fassungen von drei Instrumenten zur Erfassung von Implementierungsergebnissen. Global Implementation Research and Applications. 2021;1:183–194. Titel ins Deutsche übertragen. DOI: 10.1007/s43477-021-00019-y. [Quelle öffnen](#)
29. Verordnung (EU) 2017/745. Medizinprodukteverordnung; insbesondere Erwägungsgrund 19, Artikel 2, Artikel 5 Abs. 5 und Anhang VIII Regel 11 zur Einordnung und Klassifizierung von Software. [Quelle öffnen](#)
30. Datenschutzkonferenz. Künstliche Intelligenz und Datenschutz. Orientierungshilfe zur Auswahl, Einführung und Nutzung von KI-Anwendungen. Fassung vom 6. Mai 2024. [Quelle öffnen](#)