

Die Aufgabe der künstlichen Intelligenz endet nicht am Arztbrief

Programme hören das Patientengespräch mit und entwerfen den Arztbrief. Eine große Studie misst, wie viel Zeit das spart. Der Brief ist erst der Anfang dessen, was die künstliche Intelligenz in der ärztlichen Dokumentation übernehmen wird.

ZUSAMMENFASSUNG

Hintergrund. Ambient-Scribes hören das Patientengespräch mit und erzeugen einen Entwurf des Arztbriefs. Nach einer Umfrage der American Medical Association nutzen 81 % der US-Ärzte künstliche Intelligenz im Beruf, die klinische Dokumentation zählt zu den häufigsten Anwendungen.

Evidenz. Die erste große Multisite-Studie beobachtete 8.581 ambulant tätige Behandler an fünf US-Universitätskliniken und fand 16 Minuten weniger Dokumentationszeit pro Tag, einen Rückgang um 10 %, dazu 13,4 Minuten weniger Gesamtzeit in der elektronischen Akte und einen halben zusätzlichen Patientenbesuch je Woche. Die Zeit nach der Sprechstunde sank nicht.

Offen. Die Studie ist eine Beobachtung, die Behandler wählten das Werkzeug selbst, 79 % der Berechtigten verzichteten darauf. Die Qualität des erzeugten Briefs blieb unvermessen.

Schlussfolgerung. Belegt ist ein bescheidener Zeitgewinn bei freiwilliger Nutzung und ärztlicher Durchsicht. Der Kern der Briefarbeit, die Synthese der gesamten Krankengeschichte, liegt außerhalb der Reichweite der heutigen Systeme.

Dr. med. Christian Mauch · Facharzt für Orthopädie und Unfallchirurgie · Lesezeit 8 Minuten

Ein Ambient-Scribe ist ein Programm, das das gesamte Patientengespräch mithört und daraus einen strukturierten Dokumentationsentwurf erzeugt. Das System transkribiert nicht wörtlich. Es wählt aus, was in die Dokumentation gehört, ordnet das Gesagte in Anamnese, Befund, Beurteilung und Procedere und formuliert den Text aus. Nach einer Umfrage der American Medical Association nutzen 81 % der US-Ärzte inzwischen KI im Beruf, gegenüber 38 % im Jahr 2023. Die klinische Dokumentation zählt neben der Literaturzusammenfassung zu den häufigsten Anwendungen.³ Für die meisten

Ärzte ist der Scribe damit der erste regelmäßige Berührungspunkt mit klinischer KI.

Die bislang belastbarste Messung liefert die Arbeitsgruppe um Lisa Rotenstein, veröffentlicht im April 2026 in JAMA.¹ Die Studie beobachtete 8.581 ambulant tätige Behandler an fünf US-Universitätskliniken über gut zwei Jahre, von Juni 2023 bis August 2025. 1.809 von ihnen führten einen von drei kommerziellen Scribes ein, die übrigen 6.772 dienten als Vergleichsgruppe an denselben Häusern. Ausgewertet wurde mit einem Differenz-in-Differenzen-Ansatz, der die Veränderung bei den Nutzern gegen die der Nichtnutzer rechnet und so allgemeine Zeittrends herausfiltert. Das Ergebnis: 16 Minuten weniger Dokumentationszeit pro Tag, ein relativer Rückgang um 10 %, und 13,4 Minuten weniger Gesamtzeit in der elektronischen Akte, ein Rückgang um 3 %. Dazu kam ein halber zusätzlicher Patientenbesuch je Woche. Die Zeit, die Behandler außerhalb der Sprechstunde in der Akte verbrachten, änderte sich nicht.

Aussagekräftig an dieser Arbeit sind die Breite und der Vergleichsansatz: fünf Standorte, drei Systeme, gepoolte Daten, die weniger von den Eigenheiten einer einzelnen Einführung abhängen als die früheren Ein-Haus-Studien. Ihre Schwäche liegt im Design. Es ist eine Beobachtungsstudie. Die Behandler entschieden selbst, ob sie den Scribe nutzen. Der Ansatz filtert Zeittrends, aber nicht den Unterschied zwischen Menschen, die ein neues Werkzeug ergreifen, und Menschen, die es liegen lassen. Randomisierte Evidenz existiert bislang nur in kleinerem Maßstab. Ein in NEJM AI veröffentlichter Versuch an 238 Ärzten aus 14 Fachrichtungen fand am primären Endpunkt, der Schreibzeit je Brief, nur geringe Effekte: Mit einem der beiden geprüften Produkte sank sie um 9,5 %, mit dem anderen blieb sie unverändert. Deutlich besserten sich die sekundären Endpunkte Burnout und empfundene Arbeitslast.⁵

Was die 16 Minuten bedeuten und was nicht

16 Minuten pro Tag sind bei einem vollen Sprechstundentag ein spürbarer, aber kein umwälzender Gewinn. Die relativen Werte von 10 und 3 % zeigen, dass der weitaus größte Teil der Bildschirmarbeit bleibt. In dieselbe Rechnung gehört die Zeit außerhalb der Sprechstunde: Die Arbeit am Abend und am Wochenende, in der US-Literatur als eigene Belastungsgröße geführt, sank nicht. Der Scribe verkürzt die Dokumentation während des Arbeitstags und lässt die Arbeit danach unberührt.

Endpunkt	Veränderung	relativ
Dokumentationszeit pro Tag	–16,0 Min	–10 %
Zeit in der elektronischen Akte	–13,4 Min	–3 %

Endpunkt	Veränderung	relativ
Patientenbesuche pro Woche	+0,5	—
Zeit nach der Sprechstunde	unverändert	—

Quelle: Rotenstein LS et al., JAMA 2026. Differenz-in-Differenzen-Schätzung.

Wichtiger als die Größe des gemessenen Effekts ist, was die Studie nicht misst. Erfasst wurde die Zeit am Bildschirm, nicht die Qualität des erzeugten Dokuments. Der Arztbrief trägt die Weiterbehandlung und ist zugleich die Visitenkarte des Arztes beim nachbehandelnden Kollegen: An Vollständigkeit und Präzision liest der Empfänger ab, wie sorgfältig gearbeitet wurde. Ob ein vom Scribe entworfener und ärztlich durchgesehener Brief diesem Anspruch näherkommt als ein unter Zeitdruck selbst getippter, bleibt in dieser Arbeit ohne Antwort. Dort liegt der klinisch tragende Endpunkt, denn die Qualität der Kommunikation zwischen den Behandlern entscheidet über Doppeluntersuchungen, Medikationsfehler und Reibungsverluste. Ein begleitender Kommentar in derselben JAMA-Ausgabe benennt die Lücke grundsätzlicher: Die Forschung messe Minuten, Klicks und Codes und bleibe schwach darin, die Wirkung auf Patientenerfahrung und Versorgungsergebnisse zu erfassen.²

Die Arbeit, die der Brief nur abbildet

Hinter dieser Messlücke liegt eine grundsätzliche Grenze der heutigen Systeme. Ein tragfähiger Arztbrief, erst recht ein Gutachten, entsteht in der Sichtung vor dem Formulieren: Die gesamte Vorgeschichte des Patienten muss erfasst sein, Fremdbefunde, Bildgebung und Laborwerte müssen zusammengeführt und gewichtet werden, das Ergebnis muss der jeweiligen Gattung genügen, denn ein Gutachten folgt anderen Regeln als ein Entlassbrief, und die Beurteilung samt Therapieempfehlung muss den Leitlinien standhalten. Der heutige Ambient-Scribe leistet von alledem nichts. Er hört das Gespräch mit, auf die Akte mit ihren Vorbefunden greift er nicht zu. Seine Reichweite endet, wo die Arbeit am Brief beginnt. Die gemessenen 16 Minuten beziffern folgerichtig den kleinsten Teil der Aufgabe: das Verschriftlichen dessen, was im Behandlungszimmer gesagt wurde.

Die Selbstauswahl als Vorbehalt

Eine Zahl der Studie wird in der Rezeption selten genannt: 79 % der berechtigten Behandler verzichteten auf den Scribe. Die gemessenen Effekte beschreiben also eine Minderheit, die sich das Werkzeug selbst ausgesucht hat, und innerhalb dieser Minderheit fielen die Gewinne dort am größten aus, wo das System häufig genutzt wurde, in der hausärztlichen Versorgung stärker als in den spezialisierten Fächern.¹ Für die Übertragung auf eine Praxis heißt das: Die 16 Minuten sind kein Erwartungswert für jede Ärztin und jeden Arzt,

sie sind der Durchschnitt motivierter Anwender. Wer den Scribe flüchtig einführt, ohne nach Nutzungsbereitschaft und Fachgebiet zu unterscheiden, wird den Studieneffekt voraussichtlich unterschreiten.

Die Gegenseite der Zahl ist selbst ein Befund: Vier von fünf berechtigten Behandlern an technikaffinen Universitätskliniken verzichteten auf das Werkzeug. Die Studie erklärt die Gründe nicht, denkbar sind Gewohnheit, Misstrauen gegenüber der Entwurfsqualität oder der Aufwand der Durchsicht. Solange sie unerforscht sind, bleibt offen, ob der Scribe ein Werkzeug für eine Minderheit bleibt oder zur Standardausstattung wird.

Der Evidenzrahmen: Verbreitung vor Beleg

Die Verbreitung der Systeme läuft der Evidenz voraus. Ein Editorial in Nature Medicine vom April 2026 hält fest, dass Aussagen über den klinischen Nutzen medizinischer KI in Publikationen und Produktunterlagen häufiger werden, ohne dass Einigkeit besteht, welcher Beleg dafür genügt.⁴ Für die Praxis lässt sich das in eine Prüfliste vor der Einführung übersetzen. Erstens: Zahlen aus dem realen Betrieb verlangen, nicht aus der Produktbroschüre, für ein Umfeld, das der eigenen Praxis ähnelt. Zweitens: den Endpunkt benennen, Dokumentationszeit, Briefqualität oder Patientenzuwendung, und nach drei Monaten dagegen prüfen. Drittens: die Abbruchbedingung vorab festlegen. Ein Werkzeug, dessen Entwürfe regelmäßig nachgebessert werden müssen, verschiebt die Arbeitszeit nur an eine andere Stelle des Tages. Ein Diskussionspapier der Weltgesundheitsorganisation vom Juni 2026 zieht für die Politikgestaltung denselben Schluss: Die Systeme sollen Datenmengen erschließen und Szenarien rechnen, das menschliche Abwägen stützen und nicht ersetzen.⁶

Vor dem Einsatz: Einwilligung, Zulassung, Haftung

Für die Praxis wird die Grenze der Systeme an einem Arbeitsschritt konkret: der Freigabe. Weil das System auswählt und formuliert, entscheidet es im Entwurf, was wichtig ist. Was es weglässt oder falsch zuordnet, fällt nur auf, wenn ein Arzt den Entwurf tatsächlich liest. Die Freigabe ist deshalb der qualitätssichernde ärztliche Schritt. Sie kostet Zeit, die von den gemessenen 16 Minuten abzuziehen ist. Eine Organisation, die für die Durchsicht Zeit, Information und Zuständigkeit vorsieht, behält einen kleineren, aber belastbaren Gewinn. Wird die Durchsicht zum Klick am Ende, wandelt sich der Zeitgewinn in ein Dokumentationsrisiko.

Rechtlich sind vier Ebenen zu unterscheiden. Die erste betrifft das Gespräch selbst. Ein Scribe zeichnet das nichtöffentlich gesprochene Wort von Arzt und Patient auf. Ohne Einwilligung der Beteiligten ist das strafbar (§ 201 StGB).⁹ Die Gesprächsinhalte sind zugleich Gesundheitsdaten nach Artikel 9 der Datenschutz-Grundverordnung, deren Verarbeitung eine der eng gefassten Ausnahmen dieser Vorschrift verlangt.¹⁰ Hinzu kommt die ärztliche

Schweigepflicht nach § 203 StGB, die auch bei der Einbindung eines externen Anbieters gilt.⁹ Praktisch heißt das: Der Patient wird vor dem ersten Einsatz aufgeklärt, seine Einwilligung wird dokumentiert, und mit dem Anbieter besteht ein Auftragsverarbeitungsvertrag.

Die zweite Ebene ist das Medizinprodukterecht. Ob ein Scribe unter die Verordnung (EU) 2017/745 fällt, richtet sich nach seiner vom Hersteller festgelegten Zweckbestimmung.⁸ Ein System, das ausschließlich der Dokumentation dient, fällt regelmäßig nicht darunter. Umfasst die Zweckbestimmung diagnostische oder therapeutische Unterstützung, braucht das System ein Konformitätsbewertungsverfahren und eine CE-Kennzeichnung. Die dritte Ebene ist die KI-Verordnung. Artikel 4 der Verordnung (EU) 2024/1689 verpflichtet Betreiber seit dem 2. Februar 2025, für ein ausreichendes Maß an KI-Kompetenz ihres Personals zu sorgen, mit kontext- und risikobezogenen Maßnahmen und ohne Vorgabe einer formalen Schulung.⁷ Die vierte Ebene ist die Haftung für einen fehlerhaften, KI-erzeugten Eintrag in der Akte. Noch nicht abschließend geklärt ist die Zuordnung zwischen Hersteller, Betreiber und behandelndem Arzt bei Systemen, deren Auswahlentscheidungen sich im Einzelfall schwer nachvollziehen lassen. Bis dahin gilt: Der Arzt, der freigibt, trägt die Verantwortung für den Inhalt, und seine Durchsicht muss den Charakter einer eigenständigen Prüfung haben.

Ärztliche Einordnung

Die Scribes der heutigen Generation sind Dokumentationswerkzeuge mit Sprachverständnis. Ihr Nutzen ist belegt und an die Bedingungen gebunden, unter denen er gemessen wurde: freiwillige, häufige Nutzung und eine ärztliche Durchsicht, die als eigener Arbeitsschritt organisiert ist. Die nächste Generation wird über das Gespräch hinausgreifen, die Akte einbeziehen und Vorschläge für den Behandlungsweg machen. Damit rückt die Arbeit in ihre Reichweite, die den Wert eines Briefes oder Gutachtens bestimmt, und mit ihr rücken die Fragen nach Zweckbestimmung, Zulassung und Haftung ins Zentrum, die beim reinen Dokumentationswerkzeug noch am Rand liegen. Unvermessen bleibt bis dahin der Endpunkt, der klinisch am meisten zählt: ob die erzeugten Briefe die Weiterbehandlung verbessern, und dort endet die heutige Evidenz.

Die SummitGO-Perspektive

Der Fall der Ambient-Scribes zeigt ein übertragbares Grundprinzip medizinischer KI: Der messbare Gewinn entsteht aus der Anordnung, in die die Maschine eingebettet ist. SummitGO ordnet diese Entwicklung als Herausgeber ein und vertritt die Haltung, die sich aus der Studienlage ablesen lässt: augmented intelligence, nicht artificial intelligence. Der Maßstab dafür ist die Aufgabe, die dieser Beitrag beschrieben hat: die Sichtung und Synthese der gesamten Krankengeschichte als Zuarbeit für einen Arzt, der prüft, freigibt und verantwortet.

Fazit

Belegt ist, dass Ambient-Scribes bei freiwilliger Nutzung die tägliche Dokumentationszeit um 16 Minuten senken, bei einem halben zusätzlichen Patientenbesuch je Woche und unveränderter Arbeit außerhalb der Sprechstunde. Belegt ist ebenso, dass diese Zahlen aus einer Beobachtung selbstselektierter Nutzer stammen und die Qualität der erzeugten Dokumente nicht erfassen. Vor jedem Einsatz stehen die Aufklärung des Patienten, seine Einwilligung in die Aufzeichnung des Gesprächs und ein Auftragsverarbeitungsvertrag mit dem Anbieter. Offen bleiben die Wirkung auf die Weiterbehandlung, die Gründe der hohen Nichtnutzung und die haftungsrechtliche Zuordnung bei fehlerhaften Einträgen. Die Studienlage stützt den Einsatz als ärztlich freigegebenes Dokumentationswerkzeug. Über die Systeme der nächsten Generation, die Sichtung und Synthese der Krankengeschichte übernehmen sollen, sagt sie noch nichts.

QUELLEN

1. Rotenstein LS, Holmgren AJ, Thombley R, et al. Changes in Clinician Time Expenditure and Visit Quantity With Adoption of Artificial Intelligence-Powered Scribes: A Multisite Study. JAMA. 2026;335(16):1408–1417. doi.org/10.1001/jama.2026.2253
2. Tierney AA, Lee K, Liu VX. Ambient AI Scribes and the Quintuple Aim: What Is Counted—and What Matters. JAMA. 2026;335(16):1393–1395. doi.org/10.1001/jama.2026.3529
3. American Medical Association. 2026 Physician Survey on Augmented Intelligence. März 2026. ama-assn.org
4. Show us the evidence for the value of medical AI. Editorial. Nat Med. 2026;32:1163. doi.org/10.1038/s41591-026-04389-4
5. Lukac PJ, Turner W, Vangala S, et al. Ambient AI Scribes in Clinical Practice: A Randomized Trial. NEJM AI. 2025;2(12). doi.org/10.1056/AIoA2501000
6. World Health Organization. Artificial intelligence and evidence-informed policy: emerging challenges and opportunities. Diskussionspapier, 2. Juni 2026. who.int
7. Verordnung (EU) 2024/1689 (KI-Verordnung), Artikel 4. eur-lex.europa.eu/eli/reg/2024/1689/oj
8. Verordnung (EU) 2017/745 (Medizinprodukteverordnung). eur-lex.europa.eu/eli/reg/2017/745/oj
9. Strafgesetzbuch (StGB), § 201 Verletzung der Vertraulichkeit des Wortes und § 203 Verletzung von Privatgeheimnissen. gesetz-im-internet.de
10. Verordnung (EU) 2016/679 (Datenschutz-Grundverordnung), Artikel 9. eur-lex.europa.eu/eli/reg/2016/679/oj

Eine erste, kürzere Fassung dieses Themas erschien im Newsletter „KI und Medizin“, Ausgabe 4.

IMPRESSUM

Summitgo GmbH & Co. KG, Relenbergstraße 65, 70174 Stuttgart. Vertreten durch die Komplementärin Dr. Mauch Beteiligung GmbH (Amtsgericht Stuttgart HRB 805274), diese durch den Geschäftsführer Dr. Christian Mauch. Registergericht Amtsgericht Stuttgart, HRA 743417. USt-IdNr. gemäß § 27a UStG: [folgt nach Erteilung]. Verantwortlich für den Inhalt nach § 18 Abs. 2 MStV: Dr. Christian Mauch, Anschrift wie vor. Interessenkonflikt: SummitGO entwickelt Software zur ärztlich verantworteten KI-gestützten Dokumentation und ist damit im Gegenstand dieses Beitrags tätig. Dieser Beitrag stellt kein Produkt dar und enthält keine Leistungsbehauptung zu einem konkreten Produkt. © 2026 SummitGO. summitgo.de