

Manteltext zur Umfrage

KI-Nutzung, Kompetenz und Einstellung in der Hämatologie und Onkologie im DACH-Raum

Draft v20; Stand 10.06.2026 (angepasst an Fragebogen v50)

Friedrich Schwarz^{1,2,3}, Enver Aydilek^{1,4}, Christian Thielscher⁵, Bernd Eylert^{6,*}

1 Klinik für Hämatologie und Medizinische Onkologie, Universitätsmedizin Göttingen

2 Göttingen Campus-Institut Dynamik biologischer Netzwerke (CIDBN), Georg-August-Universität Göttingen

3 Max-Planck-Institut für Dynamik und Selbstorganisation, Göttingen

4 Klinik für Innere Medizin, Hämatologie, Onkologie, Stammzelltransplantation und Palliativmedizin, Evangelisches Klinikum Bethel, UK OWL

5 FOM Hochschule für Oekonomie & Management, Essen

6 Technische Hochschule Wildau *Studienleiter

1. Hintergrund

Künstliche Intelligenz (KI) hat in den letzten Jahren einen deutlichen Entwicklungssprung gemacht. Treiber sind insbesondere generative Large Language Models (LLMs) und darauf aufbauende Assistenz- und agentische Systeme. In der klinischen Medizin stellt sich daher zunehmend nicht mehr nur die Frage, *ob* KI-Werkzeuge Teil des Arbeitsalltags werden, sondern *wie, in welchem Umfang und unter welchen Sicherheits- und Governance-Bedingungen* sie eingesetzt werden.

Die Hämatologie und Onkologie ist ein innovations- und forschungsstarkes Fach mit datenintensiver Diagnostik, molekularen Verfahren, digitaler Pathologie, multidisziplinären Entscheidungsstrukturen und komplexer Therapiesequenzierung. Daraus ergeben sich naheliegende Anwendungsfelder für KI, zugleich aber hohe Anforderungen an Validierung, Datenschutz, Verantwortlichkeit und klinische Einbettung.

Die Dynamik der Entwicklung zeigt sich unter anderem in der AMA-Erhebung zu „Augmented Intelligence“: Der Anteil US-amerikanischer Ärzt:innen mit Kenntnis oder Nutzung von KI im klinischen Umfeld stieg von 38 % (2023) über 66 % (2024) auf 81 % (2026); 72 % gaben 2026 an, mindestens einen KI-Anwendungsfall in ihre Praxis zu integrieren¹. Recherche- und dokumentationsnahe Anwendungen werden dabei bereits am breitesten angewandt (Wissens- und Leitlinien-Zusammenfassungen ca. 39 %, dokumentationsbezogene Anwendungen ca. 28–30 %), während diagnostisch-assistive Anwendungen langsamer adaptiert werden (17 %)¹.

Für die Hämatologie und Onkologie liegen bereits fachliche Vorarbeiten vor. Mitglieder des DGHO-Arbeitskreises „Künstliche Intelligenz“ veröffentlichten 2023 eine Roadmap zu KI in der Hämatologie und Onkologie.² Das 2024 veröffentlichte DGHO-Positionspapier zu Potenzialen und Barrieren der Nutzung von Patientendaten in KI-basierter Forschung ergänzt diese Perspektive gesundheitspolitisch.³ Empirische Daten zur tatsächlichen Nutzung, Kompetenz-Selbsteinschätzung und Einstellung im DACH-Raum sind jedoch weiterhin begrenzt. Hervorzuheben sind die multidisziplinäre Befragung an deutschen Universitätskliniken⁴ oder die zweifaktorielle Einstellungsskala nach Heinrichs et al.⁵ Diese bestehenden Arbeiten adressieren hämatologisch-onkologische Besonderheiten jedoch nur sehr eingeschränkt.

Die vorliegende Umfrage soll diese Datenlücke schließen und eine Grundlage für Empfehlungen, Fortbildungsangebote und Rahmenbedingungen zum sicheren und sinnvollen KI-Einsatz in der Hämatologie und Onkologie schaffen. In die Entwicklung flossen Rückmeldungen aus DGHO-Arbeitskreisen, unter anderem Laboratorium, Präzisionsonkologie, Junge DGHO sowie Fort- und Weiterbildung, sowie Abstimmungen mit SGMO und SGH-SSH ein.

2. Motivation und Zielgrößen

2.1 Status-quo-Erhebung

Ziel ist eine DACH-spezifische Erhebung der KI-Nutzung in klinischen, wissenschaftlichen, administrativen und ausbildungsbezogenen Tätigkeitsfeldern der Hämatologie und Onkologie. Die Auswertung soll nach Position, Ausbildungsstand, Art der Hauptarbeitsstelle und Land differenziert werden.

2.2 Bedarfs- und Painpoint-Analyse

Erfasst werden Implementierungsbarrieren, verfügbare Lernzeit, bisherige Wissensquellen, Fortbildungsbedarfe, lokale Governance-Strukturen und subjektiv relevante Use-Cases. Damit soll sichtbar werden, welche praktischen Voraussetzungen aus Sicht der Anwender:innen für einen sicheren und akzeptierten KI-Einsatz fehlen.

2.3 Grundlage für fachgesellschaftliche Positionierung

Die Ergebnisse sollen als Evidenzbasis für Aus- und Weiterbildung, praxisnahe Empfehlungen und fachgesellschaftliche Positionspapiere dienen. Die Umfrage ist so angelegt, dass Wiederholungsbefragungen und damit zeitliche Trendanalysen möglich sind.

3. Methodische Eckpunkte

Design: Anonym konzipierte, freiwillige Querschnittserhebung über LimeSurvey, datenschutzkonform gehostet durch die Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen (GWDG). Nur die Einwilligung ist verpflichtend; alle weiteren Fragen können übersprungen werden. Fehlende Angaben werden als NA kodiert. Seitenbezogene Antwortzeiten werden zur Qualitätssicherung miterfasst. Die Verbreitung erfolgt über fachgesellschaftsspezifische Einladungslinks; der Einladungsweg wird als URL-Parameter ohne Personenbezug miterfasst. Die Umfrage wird auf Deutsch und Englisch angeboten; in Abstimmung mit den Schweizer Fachgesellschaften wurde auf eine französische und italienische Fassung verzichtet.

Zielgruppe: Ärzt:innen, Medizinstudierende, Mitarbeitende medizinischer Einrichtungen sowie weitere Mitglieder der beteiligten Fachgesellschaften (DGHO, OeGHO, SGMO, SGH-SSH) im DACH-Raum.

Aufbau:

- Teil A: Demografie
- Teil B: KI-Nutzung, klinische Erfahrung, Governance und Barrieren
- Teil C: Kompetenz und Ausbildung
- Teil D: Einstellungen, Chancen, Risiken und Verantwortung
- Teil E: hämatologisch-onkologischer Vertiefungsteil, überspringbar für außerhalb des Fachgebiets Tätige.

5-stufige Likert- und Likert-artige Skalen für Einstellungen, Kompetenz-Selbsteinschätzungen und Potenzialbewertungen; 6-stufige Frequenzskala für Nutzungshäufigkeit; ordinale Adoptionsskala für anwendungsbezogene Nutzung; 4-stufige Konzept-Vertrautheitsskala; 3-stufige Matrix zu Wissensquellen; Autonomiegrad-Skala mit Zusatzoption „nicht beurteilbar“; Mehrfachantworten, bei den Implementierungsbarrieren mit Begrenzung auf maximal fünf Nennungen. Die Mittelkategorie der Einstellungs-Items („neutral / weiß nicht“) fasst Indifferenz und Nichtwissen bewusst zusammen; sie wird in der Auswertung als „neutral oder unentschieden“ interpretiert. Die Opt-out-Kategorien der Matrixfragen sind konstruktsspezifisch: „keine Antwort“ (B6) für nicht zutreffende Tätigkeitsfelder, „nicht beurteilbar“ (B7, E2) bei fehlender Beurteilbarkeit sowie „irrelevant / weiß nicht“ (E1) für nicht etablierte oder nicht einschätzbare Bereiche; ihre Zuordnung zu Analysekatoren ist im Auswertungsplan festgelegt. Auf eine globale „Keine

Antwort“-Option (LimeSurvey-Einstellung shownoanswer) wurde auf Basis des kognitiven Pretests bewusst verzichtet; alle Fragen außer der Einwilligung bleiben gleichwohl überspringbar.

Bearbeitungszeit: ca. 15–20 Minuten.

Datenschutz: Im Umfragedatensatz werden keine personenbezogenen Daten gespeichert, d. h. keine Namen, E-Mail-Adressen oder IP-Adressen. Optional kann ein selbst generierter 7-Zeichen-Code für spätere Wiederholungsbefragungen angegeben werden. Ein theoretisches Re-Identifikationsrisiko in sehr kleinen Subgruppen kann nicht vollständig ausgeschlossen werden; die Auswertung erfolgt aggregiert. Die Teilnehmenden werden um einmalige Teilnahme gebeten; ein Browser-Cookie erschwert versehentliche Mehrfachteilnahmen. Kontaktaufnahme erfolgt ausschließlich außerhalb der Umfrage und auf Initiative der Teilnehmenden.

Pretest: Kognitiver Pretest mit $n \geq 5$ vor Feldstart, einschließlich Think-Aloud-Elementen und seitenweiser Zeitmessung über das LimeSurvey-Timing-Plugin.

Ethikvotum: Unbedenklichkeitsentscheidung der Ethikkommission an der Technischen Hochschule Wildau vom 23.04.2026. Für Österreich und die Schweiz wurde mit den beteiligten Fachgesellschaften abgeklärt, dass die anonyme Befragung von Fachpersonen ohne Erhebung von Patientendaten dort keiner zusätzlichen Votumspflicht unterliegt (insbesondere kein Anwendungsfall des schweizerischen Humanforschungsgesetzes).

Reporting: CROSS 2021 als primäre Reporting-Guideline für eine mögliche Publikation, ergänzt durch relevante CHERRIES-Web-Items.^{6,7}

4. Kerngedanken des Studiendesigns

4.1 Modularer Aufbau

Die Umfrage kombiniert einen allgemeinen Kern mit einem optionalen hämatologisch-onkologischen Vertiefungsteil. Der allgemeine Kern ist so formuliert, dass er grundsätzlich auch auf andere medizinische Fachgebiete übertragbar wäre. Die Blöcke A–E bezeichnen die analytische Struktur und die Fragecodes; die Anzeigereihenfolge im Online-Fragebogen ist davon abweichend drop-off-optimiert (Teile 1–7). Nicht fortlaufende Fragecodes (z. B. fehlende Codes A6 oder B5) sind Artefakte des iterativen Entwicklungsprozesses: Im Pretest oder in früheren Versionen entfernte Fragen wurden nicht neu nummeriert, damit Fragecodes über alle Versionen hinweg eindeutig referenzierbar bleiben.

- Teil A erfasst demografische und berufliche Merkmale sowie Forschungs- und Digitalmedizinbezug.
- Teil B erhebt KI-Nutzung, konkrete Anwendungsfelder, klinische Erfahrungen, Umgang mit Patient:innen-Daten, Governance, die aktive Prüfung des regulatorischen Status von KI-Medizinprodukten (B13) und Implementierungsbarrieren.
- Teil C erfasst Selbstkompetenz, verfügbare Lernzeit, Wissensquellen, Konzept-Vertrautheit und Fortbildungsbedarf.
- Teil D misst Einstellungen, Chancen- und Risikowahrnehmung, Verantwortung, Rechtssicherheit und Patient:inneninformation.
- Teil E fokussiert auf hämatologisch-onkologische Anwendungsbereiche, deren Entwicklungsstand, erwarteten Nutzen und Kontakt zu zertifizierten KI-Medizinprodukten sowie die Mitgliedschaft in den beteiligten Fachgesellschaften (A9); bei Übertragung auf andere Fachgebiete wird die Mitgliedschaftsfrage im jeweiligen Fachmodul angepasst.

4.2 Messlogik und Auswertung

Ein zentrales Designmerkmal ist die parallele Erhebung von Ist-Zustand und Soll-Zustand bzw. Erwartung. In Teil B werden für dieselben acht Tätigkeitsfelder sowohl der aktuelle bzw. geplante Nutzungsstand (B6) als auch die gewünschte zukünftige Rolle der KI (B7) abgefragt.

Kompetenz wird als Selbstauskunft erhoben und bewusst nicht als Wissenstest operationalisiert, um Bearbeitungszeit und Akzeptanz in der klinischen Zielgruppe zu wahren. Plausibilitäts- und Konsistenzanalysen können Selbsteinschätzung, tatsächliche Nutzung, Wissensquellen und fachspezifische Angaben miteinander vergleichen.

4.3 Regulatorische Abgrenzung

Die Erhebung unterscheidet durchgängig zwischen allgemeinen KI-Werkzeugen bzw. LLMs (z. B. ChatGPT, Claude, Gemini, Copilot) und zertifizierten KI-Medizinprodukten nach MDR/IVDR bzw. MepV/IvDV. Diese Abgrenzung ist zentral, weil Zweckbestimmung, Validierungsanforderungen, institutionelle Freigabe und Verantwortlichkeiten unterschiedlich sind. Entsprechend werden sowohl formale Freigaben oder Policies auf Ebene des Arbeitsbereichs als auch die aktive Prüfung des regulatorischen Status von KI-Medizinprodukten erhoben.

4.4 Anonymität, longitudinale Verknüpfung und Qualitätssicherung

Die Befragung erhebt keine direkten Identifikatoren und ist als anonyme Online-Erhebung mit optionaler pseudonymer Längsschnittverknüpfung konzipiert. Für spätere Wiederholungsbefragungen kann freiwillig ein 7-Zeichen umfassender Self-Generated Identification Code (SGIC) angegeben werden. Der Code basiert auf stabilen, von den Teilnehmenden selbst generierten Merkmalen und orientiert sich an etablierten Verfahren für die Verknüpfung anonymer bzw. pseudonymer Längsschnittdaten.^{8,9} Da die Kombination aus Code und demografischen Angaben in sehr kleinen Subgruppen theoretisch re-identifizierend wirken kann, wird dieses Restrisiko in den Umfragehinweisen transparent kommuniziert. In Auswertung und Berichterstattung wird das verbleibende Restrisiko zusätzlich durch Zellenunterdrückung bei $n < 5$ in Untergruppentabellen abgefangen.

Die Qualitätssicherung erfolgt über Pretest, technische Prüfung des LimeSurvey-Instruments und post-hoc-Analyse der Antwortzeiten. Auf Instructed-Response- oder Bogus-Items wird verzichtet, um die Bearbeitungszeit gering zu halten und die Akzeptanz in der Zielgruppe nicht zu beeinträchtigen.

5. Limitationen

Die Untersuchung ist als freiwillige Online-Querschnittserhebung konzipiert. Sie erlaubt daher keine Kausalaussagen und ist anfällig für Selbstselektions- und Non-Response-Bias; insbesondere KI-affine Personen könnten überrepräsentiert sein. Repräsentativität im DACH-Raum hängt von Reichweite, Teilnahmequote und Zusammensetzung der beteiligten Fachgesellschaften ab. Die Mitgliederbasis der beteiligten Fachgesellschaften ist nicht zwingend repräsentativ für alle hämatologisch-onkologisch tätigen Ärzt:innen. Als offene, anonyme Umfrage ohne Zugangstokens kann eine Mehrfachteilnahme trotz Cookie-basierter Erschwerung und der Bitte um einmalige Teilnahme technisch nicht vollständig ausgeschlossen werden.

Nutzung, Kompetenz und potentiell kritische KI-Erfahrungen werden als Selbstbericht erhoben. Dadurch sind Über- und Unterschätzungen, Erinnerungsverzerrungen und Social-Desirability-Effekte möglich. Besonders Angaben zu fehlerhaften oder potenziell schädlichen KI-Empfehlungen können trotz Anonymität unterberichtet werden.

Der hämatologisch-onkologische Vertiefungsteil (Teil E) kann zusätzliche Selbstselektion erzeugen. Die Begrenzung der Implementierungsbarrieren auf maximal fünf Nennungen verbessert die Priorisierung, kann seltenere Barrieren aber untererfassen.

Adoptionsstadium und gewünschter Autonomiegrad messen unterschiedliche Konstrukte; die Gap-Analyse wird deshalb nicht als direkte Differenz ordinaler Skalen interpretiert. Die Mittelkategorie „neutral / weiß nicht“ der Einstellungsskalen trennt nicht zwischen neutraler Haltung und Nichtwissen; Anteile der Mittelkategorie werden daher nicht als ausschließlich neutrale Positionen interpretiert.

Trotz des optionalen SGICs und der Zellenunterdrückung bei $n < 5$ bleibt bei sehr kleinen demografischen Subgruppen ein theoretisches Re-Identifikationsrisiko bestehen.

Referenzen

1. American Medical Association. *2026 Physician Survey on Augmented Intelligence*. <https://www.ama-assn.org/practice-management/digital-health/2026-physician-survey-augmented-intelligence> (2026).
2. Rösler, W. *et al.* An overview and a roadmap for artificial intelligence in hematology and oncology. *J. Cancer Res. Clin. Oncol.* **149**, 7997–8006 (2023).
3. *Potenzial Und Barrieren Bei Der Nutzung von Patientendaten in KI-Basierter Forschung Am Beispiel Der Hämatologie Und Medizinischen Onkologie*. vol. 21 (DGHO, Berlin, 2024).
4. Maassen, O. *et al.* Future medical artificial intelligence application requirements and expectations of physicians in german university hospitals: Web-based survey. *J. Med. Internet Res.* **23**, e26646 (2021).
5. Heinrichs, H., Kies, A., Nagel, S. K. & Kiessling, F. Physicians' attitudes toward artificial intelligence in medicine: Mixed methods survey and interview study. *J. Med. Internet Res.* **27**, e74187 (2025).
6. Sharma, A. *et al.* A consensus-based checklist for reporting of survey studies (CROSS). *J. Gen. Intern. Med.* **36**, 3179–3187 (2021).
7. Eysenbach, G. Improving the quality of web surveys: the checklist for reporting results of internet E-surveys (CHERRIES). *J. Med. Internet Res.* **6**, e34 (2004).
8. Schnell, R., Bachteler, T. & Reiher, J. Improving the use of self-generated identification codes. *Eval. Rev.* **34**, 391–418 (2010).
9. Audette, L. M., Hammond, M. S. & Rochester, N. K. Methodological Issues With Coding Participants in Anonymous Psychological Longitudinal Studies. *Educ. Psychol. Meas.* **80**, 163–185 (2020).