



Praxisrelevante KI-Meilensteine in der Onkologie

AG Krebsregister
Mitgliederversammlung
70. Jahrestagung der GMDS, 7.-11.9.2025 in Jena
9.9.2025, 11:00-12:30

Prof. Dr. med. Marco Halber, Dipl.-Betriebswirt (FH)
<https://www.krebsregister-bw.de>, <https://halber.de>

Nützliche KI in der Onkologie: Fünf Beispiele

MASAI, MSK, TrialGPT, TrialMatchAI, MIKA

Screening

- **Mammographie** (MASAI-Studie)
...steigert Entdeckungsrate und diagnostische Genauigkeit

Diagnose

- **Digitale Pathologie** (MSK-CC)
...priorisiert Befunde und erhöht diagnostische Genauigkeit

Studien

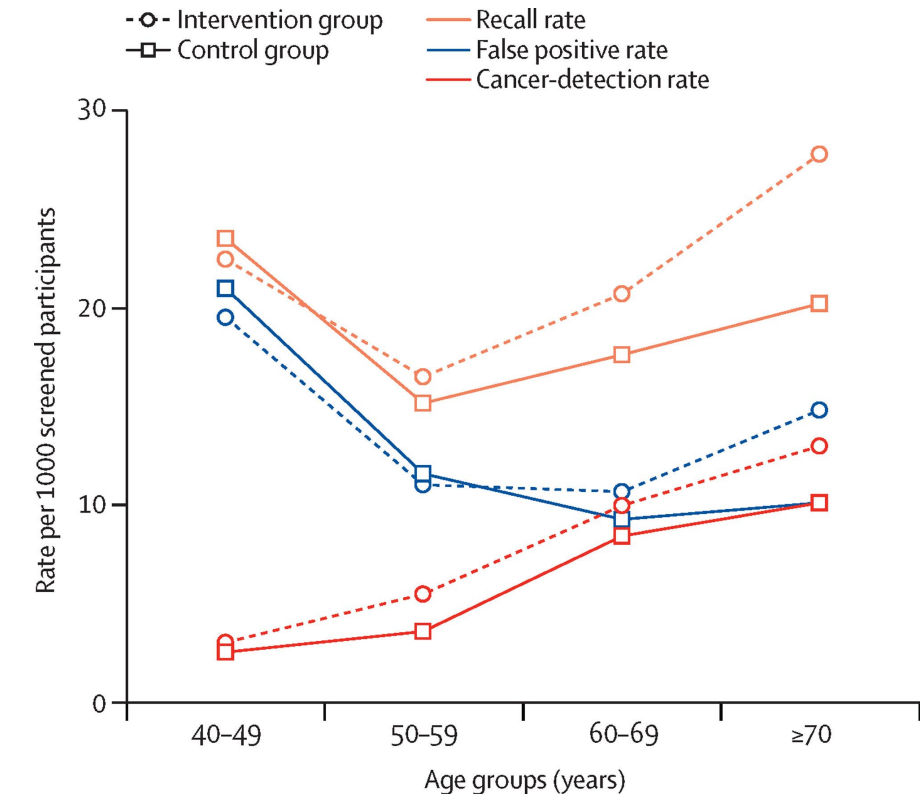
- **Studienrekrutierung** (TrialGPT und TrialMatchAI)
...erhöht Genauigkeit und Effizienz

Outcome

- **Patientenunterstützung** (MIKA)
...erweitert die Teilhabe Betroffener, bietet Orientierung über Therapien, stärkt Selbstmanagement, reduziert mentale Belastung etc.

Mammographie (MASAI-Studie)

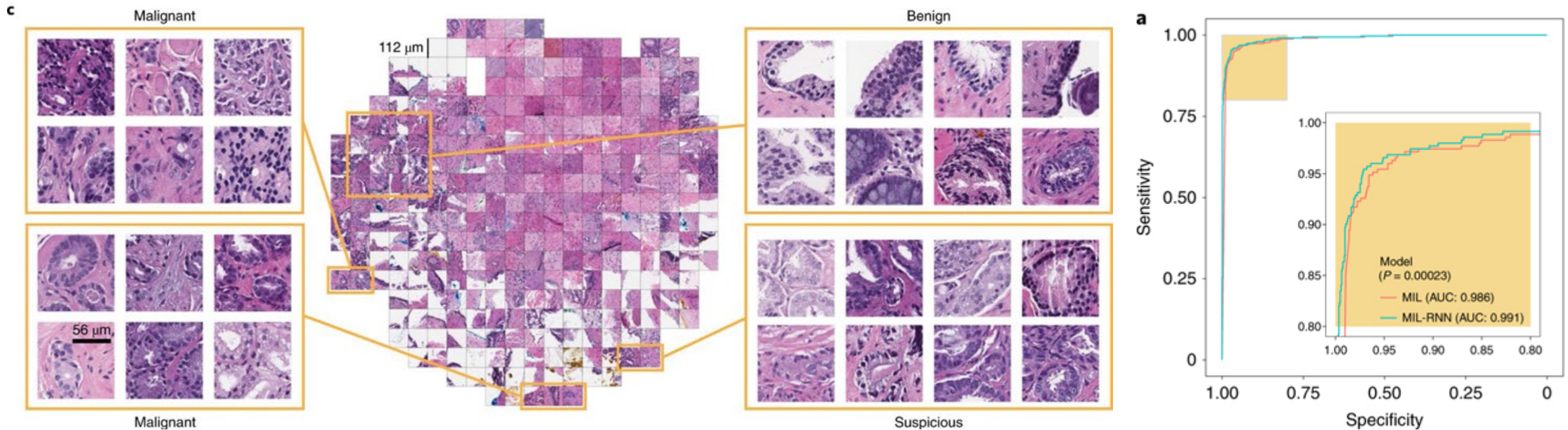
	Intervention group (n=53 043)	Control group (n=52 872)	Proportion ratio	p value
Early screening performance				
Number of recalls	1110	1027	..	..
Recall rate	2.1% (2.0–2.2)	1.9% (1.8–2.1)	1.08% (0.99–1.17)	0.084
Number of detected cancers	338	262	..	..
Cancer-detection rate, per 1000	6.4 (5.7–7.1)	5.0 (4.4–5.6)	1.29 (1.09–1.51)	0.0021
Number of false positives	772	765	..	..
False positive rate	1.5% (1.4–1.6)	1.4% (1.3–1.6)	1.01% (0.91–1.11)	0.92
Positive predictive value of recall	30.5% (27.8–33.3)	25.5% (22.9–28.3)	1.19% (1.04–1.37)	0.012
Workload				
Number of screen readings	61 248	109 692		
Number of consensus meetings	2023	1958		
Consensus meeting rate	3.8%	3.7%		



Quelle: Hernström, V. et al. Screening performance and characteristics of breast cancer detected in the Mammography Screening with Artificial Intelligence trial (MASAI): a randomised, controlled, parallel-group, non-inferiority, single-blinded, screening accuracy study. The Lancet Digital Health 7 (3), e175 - e183 (3/2025) - [https://www.thelancet.com/journals/landig/article/PIIS2589-7500\(24\)00267-X/fulltext](https://www.thelancet.com/journals/landig/article/PIIS2589-7500(24)00267-X/fulltext)

Digitale Pathologie – am MSKCC schon seit 2019

Deep-Learning-Modell für digitale Gewebeschnitte mit weakly supervised learning (fast ohne Annotation):
AUC > 0,98 mit 100 % Sensitivität; > 75 % der Gewebeschnitte können sicher übergangen werden
(44.732 Gewebeschnitte, 15.187 Fälle) – zahlreiche Weiterentwicklungen seitdem; Beispiel Prostata:

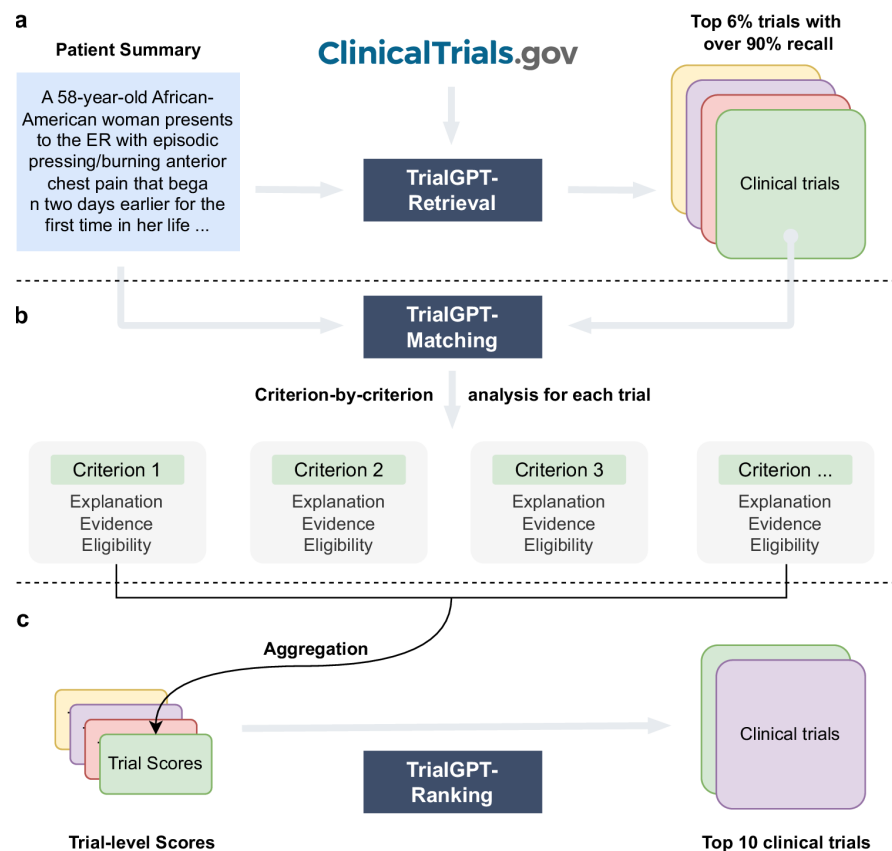


Quelle: Campanella, G. et al. **Clinical-grade computational pathology using weakly supervised deep learning on whole slide images.**
Nat Med 25, 1301–1309 (2019). <https://doi.org/10.1038/s41591-019-0508-1>

Patient-Trial-Matching (TrialGPT)

Stand: 18.1.2024

- Genauigkeit (Accuracy): 87,3 % (183 synth. Patienten)
- Richtig-Positiv-Rate (Recall) der Studien: > 90 %
- Verringerung der Screening-Zeit um 43 %



c

		Annotator A Acc: 0.904 (0.876-0.935)				Annotator B Acc: 0.876 (0.843-0.913)				Annotator C Acc: 0.916 (0.888-0.947)				TrialGPT Acc: 0.899 (0.868-0.930)			
Ground-truth	Included	0.96	0.01	0.03	0.00	0.94	0.01	0.05	0.00	0.95	0.01	0.04	0.00	0.90	0.01	0.09	0.00
	Not Incl.	0.04	0.91	0.04	0.00	0.13	0.48	0.30	0.09	0.00	0.91	0.09	0.00	0.09	0.78	0.13	0.00
	No Info.	0.04	0.07	0.86	0.03	0.09	0.03	0.87	0.02	0.06	0.04	0.89	0.01	0.01	0.04	0.91	0.04
	Not Appl.	0.00	0.00	0.11	0.89	0.00	0.00	0.00	1.00	0.00	0.00	0.11	0.89	0.00	0.00	0.00	1.00
		Included	Not Incl.	No Info.	Not Appl.	Included	Not Incl.	No Info.	Not Appl.	Included	Not Incl.	No Info.	Not Appl.	Included	Not Incl.	No Info.	Not Appl.

d

		Annotator A Acc: 0.898 (0.876-0.921)				Annotator B Acc: 0.894 (0.869-0.918)				Annotator C Acc: 0.891 (0.868-0.915)				TrialGPT Acc: 0.859 (0.833-0.885)			
Ground-truth	Excluded	0.90	0.10	0.00	0.00	0.90	0.00	0.10	0.00	0.90	0.00	0.10	0.00	1.00	0.00	0.00	0.00
	Not Excl.	0.00	0.93	0.02	0.04	0.01	0.90	0.08	0.01	0.01	0.90	0.05	0.04	0.01	0.86	0.04	0.10
	No Info.	0.01	0.17	0.78	0.04	0.01	0.06	0.89	0.04	0.01	0.13	0.86	0.00	0.01	0.09	0.82	0.08
	Not Appl.	0.00	0.13	0.00	0.87	0.00	0.13	0.07	0.80	0.00	0.11	0.04	0.85	0.00	0.02	0.00	0.98
		Excluded	Not Excl.	No Info.	Not Appl.	Excluded	Not Excl.	No Info.	Not Appl.	Excluded	Not Excl.	No Info.	Not Appl.	Excluded	Not Excl.	No Info.	Not Appl.

Predictions

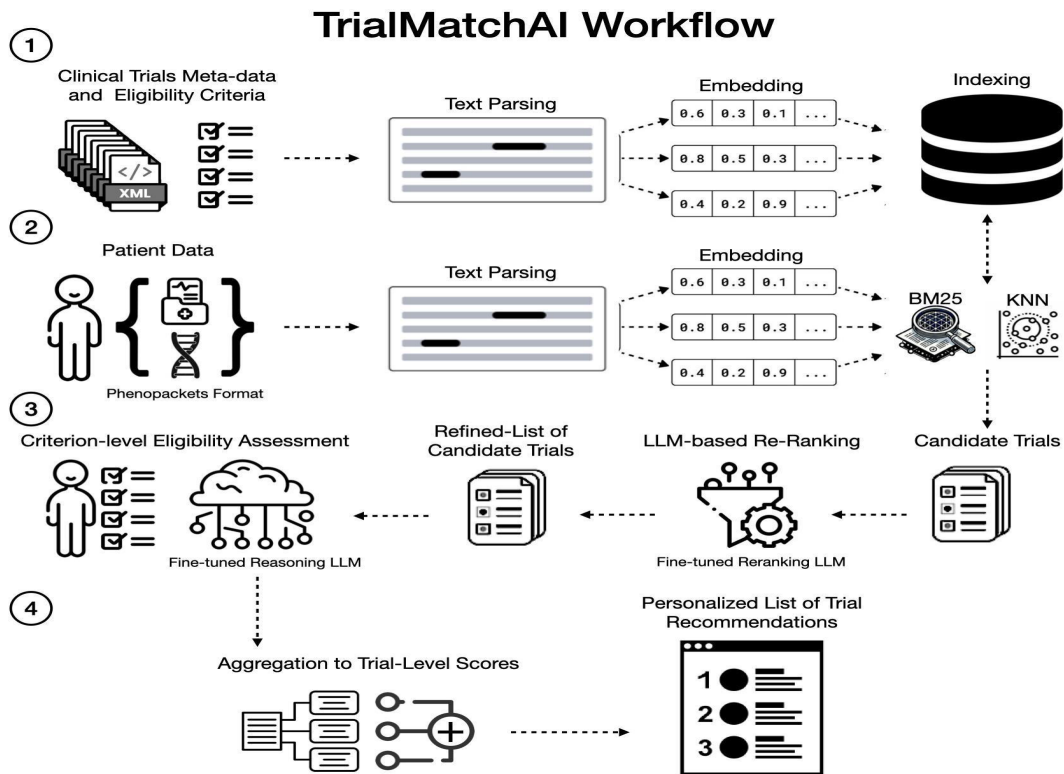
Quelle: Jin, Q., Wang, Z., Floudas, C.S. et al. Matching patients to clinical trials with large language models. Nat Commun 15, 9074 (2024). <https://doi.org/10.1038/s41467-024-53081-z>

Trial-Matching für die Onkologie (TrialMatchAI)

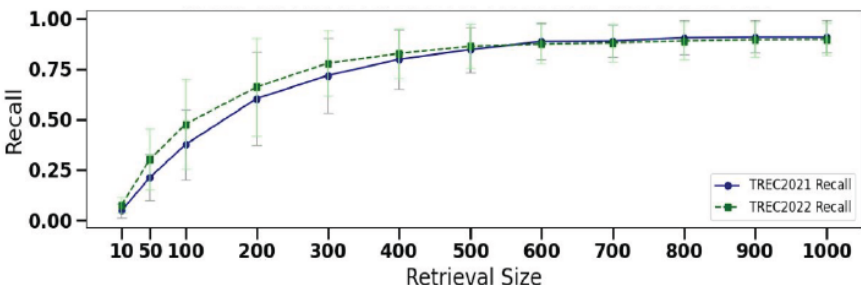
Stand: 13.5.2025



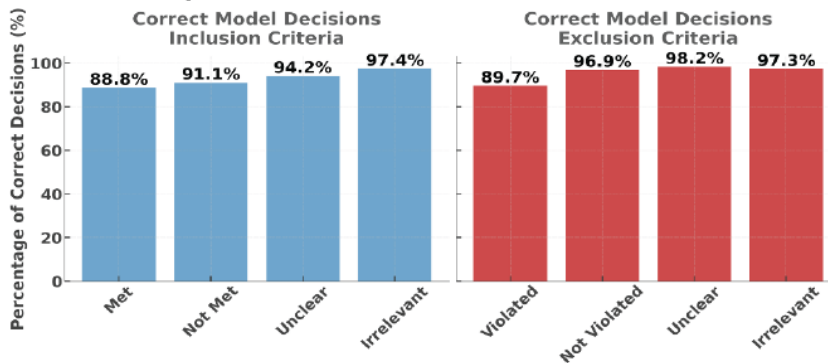
- Basis: Mehrere Open-Source-RAG-LLM und ML-Tools
- Nutzt strukturierte Daten und unstrukturierte Arzt-Notizen sowie Pathologie-Befunde etc.
- **Trefferquote: 92 %** (mindestens eine passende Studie unter den Top-20 (85 % unter den Top-5))
- Genauigkeit (Accuracy): > 90 % (vor allem Biomarker-basiert)



A Recall Performance Results For a Range of Retrieval Depths



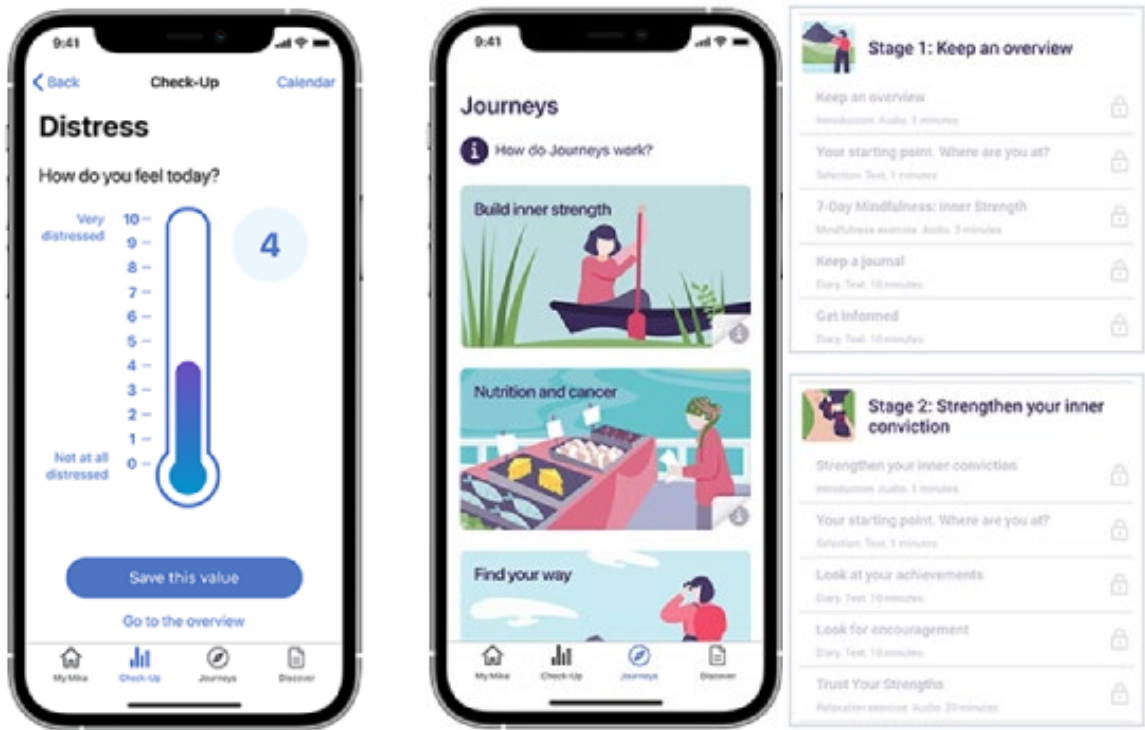
B Expert Evaluation Results on Patient-Criterion Pairs



Quelle: Abdallah, M. et al: TrialMatchAI: An End-to-End AI-powered Clinical Trial Recommendation System to Streamline Patient-to-Trial Matching. arXiv:2505.08508 [cs.AI] (5/2025), <https://doi.org/10.48550/arXiv.2505.08508>

Patientenunterstützung (MIKA-App)

- AI-gestützte Patienten-App, Klasse-IIa-Medizinprodukt (EU MDR)
- Erfasst Symptome, Nebenwirkungen, Behandlungsfortschritt (PROM)
- Datenlieferung anonymisiert an (auch industrielle) Kooperationspartner
- Erinnerung an **Medikamente** und **Termine**
- „Journeys“: Verhaltenstherapie gegen Angst, Stress und Erschöpfung



Dr. G. Finke Dr. J.S. Raue
Fosanis GmbH, Berlin (2017)

„...kostenlos. Für immer.
Wir nehmen kein Geld von Menschen,
die von Krebs betroffen sind.“

	Intervention		Control		Analysis of covariance		
	n	mean (SD)	n	mean (SD)	F test (df)	P val	ηp²
Intention-to-treat							
Baseline	99	6.0 (2.1)	119	6.0 (2.4)			
Follow-up	99	5.1 (2.5)	119	5.9 (2.5)			
Change	99	-0.8 (2.8)	119	-0.1 (2.9)	4.7 (1, 215)	.03	0.02
Per-protocol							
Baseline	50	6.0 (2.2)	74	5.7 (2.2)			
Follow-up	50	4.8 (2.8)	74	5.9 (2.4)			
Change	50	-1.2 (3.2)	74	0.2 (2.8)	6.9 (1, 121)	.01	0.05

Quelle: <https://de.mika.health/> , Springer. F. et al.: Digital Therapeutic (Mika) Targeting Distress in Patients With Cancer: Results From a Nationwide Waitlist Randomized Controlled Trial. J Med Internet Res 2024;26:e51949, <https://www.jmir.org/2024/1/e51949>

Lesetipp zu Medizin-KI-Technikfolgen

Goldschmidt/Deserno/
Winter (Hrsg.)

gmds Deutsche Gesellschaft für
Medizinische Informatik,
Biometrie und
Epidemiologie e.V.

KI in der Medizin

Folgenabschätzung für die Forschung und Praxis



Gesundheitswesen in der Praxis

 medhochzwei

- **Überblick & Ethik**
Chancen, Risiken und ethische Fragen zu Verantwortung, Sicherheit und Haftung; Biographische Analyse zu Siegfried Koller
- **Leitlinien & Governance**
Praktischer Handlungsrahmen durch ethische Leitlinien, Beispiel: KI-Governance UK Erlangen
- **Praktische Anwendungen**
Beispiele: Diagnostik, Kardiologie, Antibiotika-Management, klinische Kodierung
- **Rechtliche Rahmenbedingungen**
Haftung und Datenschutz nach MDR und EU AI Act (dabei ärztliche Verantwortung zentral)
- **Qualität & Bias**
Gesamter Learning-Zyklus muss geprüft werden (Qualität, Nutzen und Verzerrungen)