

ABSTRACTS CONFERENCE PROCEEDINGS

2. DGDM Symposium für Digitale Medizin

23.09.25 bis 24.09.2025

Hasso-Plattner-Institut, Potsdam

HERAUSGEGEBEN VON / PUBLISHED BY

Deutsche Gesellschaft für Digitale Medizin e.V

Mühlenstraße 8a

14167 Berlin

Sowie dem aktuellen Vorstand / and the current board:

Dr. Lars Masannek

Sebastian Schulz

Dr. Konstanze Betz

Linea Schmidt

Dr. Julius Gerstmeyer

Dr. Anne-Sophie Kubasch

Dr. Stefan Krautschneider

Fabian Schmid

Und der mitausrichtenden wissenschaftlichen Beirätin / and the co-hosting scientific advisor:

Prof. Dr. Ariel Dora Stern

DOI des gesamten Bandes / DOI of the whole proceedings:

10.5281/zenodo.17143736

Zitiervorschlag / Please cite as:

Masannek, L., Schulz, S., Betz, K., Schmidt, L., Gerstmeyer, J., Kubasch, A.-S., Krautschneider, S., Schmid, F., & Stern, A. D. (2025, September 23). Conference Proceedings - 2. DGDM Symposium für Digitale Medizin. 2. DGDM Symposium für Digitale Medizin (2. DGDM), Hasso-Plattner-Institute, Potsdam. Deutsche Gesellschaft für Digitale Medizin. <https://doi.org/10.5281/zenodo.17143736>

Für jeden Beitrag finden Sie ein eigenes DOI für spezifische Beitragszitate im jeweiligen Abstract.

For each abstract you will find its own DOI for specific citations below.

Eine digitale DOI-Sammlung ist verfügbar / a digital DOI collection can be accessed at:

VORWORT / FOREWORD

German

Medizin und Technik sind längst keine getrennten Welten mehr – digitale Verfahren und KI sind heute über klinische Domänen hinweg integraler Bestandteil ärztlichen Handelns. Das zeigen die Beiträge dieses Symposiums eindrucklich: von KI-gestützter Bildgebung in der Prostata-MRT über klinische Entscheidungsunterstützung mit Millionen von Routinedaten bis hin zu Wearables, die bei Multipler Sklerose feinmotorische Einschränkungen oder bei chronischen Neuropathien die Schlafqualität messen.

Es ist nicht selbstverständlich, dass eine junge Fachgesellschaft wie die DGDM in so kurzer Zeit wächst – und dass die Einreichungen in diesem Jahr noch vielfältiger und qualitativ stärker sind als im ersten. Darum haben wir uns entschieden, erstmals einen Abstractband herauszugeben, um diese Breite und Tiefe Digitaler Medizin auch über das Symposium hinaus sichtbar zu machen.

Das Symposium und dieser Band sind das Ergebnis rein ehrenamtlicher Arbeit – getragen von einem Funken, der überspringt, motiviert und interdisziplinär wie interprofessionell weiterträgt.

Wir freuen uns schon jetzt auf das nächste Jahr und gemeinsam mit allen Teilnehmenden die Digitale Medizin von morgen zu gestalten.

English

Medicine and technology are no longer separate realms – digital methods and AI have already become integral parts of clinical practice across the board. The contributions in this symposium demonstrate this impressively: from AI-assisted prostate MRI imaging, to clinical decision support based on millions of routine data records, to wearables capturing fine motor impairments in multiple sclerosis or sleep quality in chronic neuropathies.

It is far from self-evident that a young society like the DGDM has grown so rapidly – and even more remarkable that this year's digital medicine submissions are broader in scope and higher in quality than in our first year. This is why we have decided to publish an abstract volume for the first time, making this breadth and depth visible beyond the symposium itself.

Both the symposium and this volume are the result of entirely voluntary efforts – driven by a spark that inspires, motivates, and carries forward across disciplines and professions.

We are already looking forward to the coming year and to shaping the digital medicine of tomorrow together with all participants.

Inhaltsverzeichnis / Table of Contents

KI in der Prostata-MRT: Real-World-Daten aus der Praxis.....	8
Entwicklung und Implementierung von KI-basierter klinischer Entscheidungsunterstützung in der stationären Versorgung.....	10
Kennzahlenentwicklung zur ROI-Bewertung von KHZG-konformen Patientenportalen am Beispiel der No-Show-Rate in der Neurochirurgie	11
Tastaturnutzung als digitaler Marker feinmotorischer Funktion bei Multipler Sklerose	13
Analyse der Schlafqualität mittels klinischer und digitaler Messmethoden bei Patienten mit chronischer inflammatorischer demyelinisierender Polyneuropathie	15
Cost-effectiveness and -utility of a tele-neurologic intervention alongside a stepped wedge randomized controlled trial in a rural area within Germany (NeTKoH)	17
Digitale Navigator:innen zur Akzeptanz- und Kompetenzentwicklung mit Mental Health Apps: Ergebnisse einer Mixed-Method-Interventionsstudie (DigiNavi)	19
Interdisziplinäre Mustererkennung digitaler Phänotypen in der Multiplen Sklerose mittels maschinellen Lernens: ein Use Case aus Gang- und MRT-Daten.....	21
Safety and Security Implications of LxM-Enhanced and Agentic Systems in Healthcare ...	23
Leveraging ChatGPT for Systematic Reviews - A feasibility study and framework proposal	24
Wie sehr stimmen Large Language Modells in ihren Empfehlungen zu Diagnostik und Therapie eines*r Patient*in mit den Empfehlungen der Leitlinie überein? Eine Betrachtung am Beispiel der S2K Leitlinie zum Hämolytisch-urämischem Syndrom	26
Towards Personalized Neuroscience: Evaluating Individual-Level.....	28
Information in Neural Mass Models	28
Migraine Research with Smartwatch Application (MIRA) - A	30
prospective migraine study with Wearables	30
Bewertung der Akzeptanz von künstlicher Intelligenz in der Kinder- und Jugendmedizin – eine qualitative Fokusgruppenstudie mit Kindern und Jugendlichen	31
Commercialization of DiGA in Germany: Success Factors, Barriers, and Strategic Implications.....	31
Digital Influence: Kategorisierung deutscher Medfluencer*innen auf Instagram	34

Who to Trust: the Patient or the Data? A 3-Month Telemonitoring Study in Youths with Post-COVID Syndrome	36
Large language models for dermatological image interpretation – a comparative study ...	39
Translating Neurologist Notes into Disability Scores: Developing a Large Language Modell for Multiple Sclerosis Research and Care	41
Bedarfe und Wünsche von Hausärzten zu Informationen über digitale Gesundheitsanwendungen im Bereich Psyche	43
Regulatorische Rahmenbedingungen und Fallbeispiele für multizentrische, dezentrale komplett digitale Studien mit Medizinprodukten	45
Augmented Emotional Reality – Multimodale Emotionsübermittlung in der Telemedizin ..	49
Medical Billing Data: A hidden treasure for predicting Post-Endoscopic Mucosal Resection Adverse Events?.....	51
Acceptance of and Engagement with VitaLo	52
Kombinierte Versorgung als Konzept der digitalen Reha-Nachsorge: Ergebnisse der Patientenzufriedenheit und Benutzerfreundlichkeit	53
Digitale Transformation im Versorgungskontext: Praxisnahe Unterstützung hausärztlicher Teams durch Digitalisierungslots:innen	55
Digital Health Spark - Igniting need-driven innovation in healthcare: A study on motivation, self-confidence and understanding in technology transfer and business creation	57
Mehr Likes als Wirkung? Nutzung digitaler psychologischer Unterstützungsangebote bei Jugendlichen mit psychischen Störungen.....	59
Abstract	59
Effects of an app-based physical training for long and post COVID treatment (A PILOT study) on physical performance and psychological health.....	61
Status Quo der Digitalisierung in der Dermatologie – Ergebnisse einer nationalen Befragung	63
Digitalisierung im Sozial- und Gesundheitswesen: Zwischen Offenheit und strukturellen Hürden.....	65
Treating Metabolic Syndrome and Steatotic Liver Disease: A Systematic Literature Review of Conventional, Digital, and Personalized Just-in-Time Treatment Interventions	67
Kombinierte Versorgung als Konzept der digitalen Reha-Nachsorge:	69
Ergebnisse der Patientenzufriedenheit und Benutzerfreundlichkeit	69

Strukturierte Lehre der Grundlagen von Digitaler Medizin im Medizinstudium in Form von digitalen Lerneinheiten in Articulate Rise.....	71
Internationale Strategien zur Verbesserung der Interoperabilität im Austausch von Gesundheitsinformationen: Ein Scoping Review	72
Digitale Kluft statt digitaler Teilhabe? Sozioökonomische Ungleichheiten in der digitalen Gesundheitskompetenz.....	75
Addressing AI Bias Through Data Disclosure: Evidence from a randomized Experiment among Physicians	80

KI in der Prostata-MRT: Real-World-Daten aus der Praxis

Werner DJ^{1*}, Schalber C², Gaul S², Schöpfer K², Liepe L², Tresch A³,
Rehman A³, Rey JW⁴

*präsentierender Autor

1 davidwerner1903@gmail.com, Radiologie Rhein-Nahe, Krankenhaus St. Marienwörth, 55543 Bad Kreuznach, Rheinland-Pfalz, Deutschland

2 Radiologie Rhein-Nahe, Krankenhaus St. Marienwörth, 55543 Bad Kreuznach, Rheinland-Pfalz, Deutschland

3 Institut für Medizinische Statistik und Bioinformatik, Medizinische Fakultät, Universität zu Köln, 50923 Köln, Nordrhein-Westfalen, Deutschland

4 Department of Gastroenterology and Endoscopy, Medical Center Osnabrueck, Osnabrueck 49076, Niedersachsen, Germany

DOI: 10.5281/zenodo.17052575

Einleitung: Die multiparametrische Magnetresonanztomographie (mpMRT) ist essenziell für die Detektion des Prostatakarzinoms, jedoch zeitaufwendig und interindividuell variabel in der Befundung. Künstliche Intelligenz (KI) bietet Potenzial zur Standardisierung und Effizienzsteigerung, doch es fehlt an belastbarer Real-World-Evidenz für ihren klinischen Nutzen. Ziel dieser Studie ist die Evaluation einer KI-gestützten Prostata-MRT-Befundung unter Alltagsbedingungen in einer radiologischen Praxis.

Methoden: In einer retrospektiven, randomisierten Studie wurden mpMRT-Daten von Patienten mit klinischem Verdacht auf Prostatakarzinom analysiert. Drei Gruppen unterschiedlich erfahrener Radiolog*Innen befundete die Bildgebung – mit und ohne KI-Unterstützung. Primäre Endpunkte waren Befundungszeit, Übereinstimmung mit klinisch-histopathologischen Ergebnissen und diagnostische Sicherheit.

Ergebnisse: Die Nutzung der KI führte zu einer signifikanten Reduktion der Befundungszeit und einer höheren diagnostischen Konsistenz zwischen Radiologen unterschiedlicher Erfahrungsstufen. Die Sensitivität und Spezifität der KI-gestützten Befunde entsprachen weitgehend den Expertenbefunden. Die Radiologen bewerteten die KI als hilfreich, insbesondere für unerfahrene Befunder.

Diskussion: Die Studie zeigt, dass KI in der Prostata-MRT die Effizienz und Standardisierung der Diagnostik in einer realen klinischen Umgebung verbessert. Dies unterstreicht das Potenzial digitaler Diagnostik mit Real-World-Daten und ermöglicht eine evidenzbasierte Implementierung von KI in die radiologische Routine. Weitere prospektive Studien sind erforderlich, um langfristige Effekte und regulatorische Anforderungen zu adressieren.

Citation:

Baehr, F., Ludolph, P., Dastur, M., Hugo, J., Schädler, D., Behn, C.-O., Wehkamp, K., & Ladenbauer, J. (2025, September 23). Entwicklung und Implementierung von KI-basierter klinischer Entscheidungsunterstützung in der stationären Versorgung. 2. DGDM Symposium für Digitale Medizin (2. DGDM), Hasso-Plattner-Institut, Potsdam.
<https://doi.org/10.5281/zenodo.17058306>

Entwicklung und Implementierung von KI-basierter klinischer Entscheidungsunterstützung in der stationären Versorgung

F. Baehr^{1*}, P. Ludolph¹, M. Dastur¹, J. Hugo¹, D. Schädler², C.-O. Behn³, K. Wehkamp⁴, J. Ladenbauer¹

*präsentierender Autor

1 f.baehr@tiplu.de, Tiplu GmbH, Deutschland

2 Klinik für Anästhesiologie und Operative Intensivmedizin, Universitätsklinikum Schleswig-Holstein, Deutschland

3 Klinik für Innere Medizin I, Universitätsklinikum Schleswig-Holstein, Deutschland

4 Department für Medizinmanagement, Medical School Hamburg, Deutschland

DOI: 10.5281/zenodo.17058306

Einleitung:

KI-basierte Früherkennung und Risikoprädiktion medizinischer Ereignisse durch Clinical Decision Support Systems (CDSS) birgt großes Potenzial für die Verbesserung der Patientenversorgung. Insbesondere für die Früherkennung von Sepsis besteht Evidenz, dass KI-basierte CDSS zu einer Mortalitätsreduktion im Krankenhaus beitragen können [1].

In diesem Beitrag stellen wir die Entwicklung, Implementierung und Evaluation eines CDSS mit Machine Learning (ML)-basierter Prädiktion in der stationären Versorgung vor.

Methoden:

Der Lösungsansatz basiert auf einer in Deutschland einzigartigen Infrastruktur für föderiertes ML, die interoperable, strukturierte und unstrukturierte klinische Routinedaten aus 10 Millionen Patientenakten von über 130 Krankenhäusern nutzt. Die Prädiktionen und generierten Erklärungen der trainierten Modelle wurden auf einem separaten Teil des Trainingsdatensatzes retrospektiv evaluiert und, nach Integration der Modelle in die CDSS-Software, von Mediziner:innen prospektiv in einem simulierten Anwendungsszenario auf Angemessenheit überprüft. Das CDSS wird im Rahmen einer klinischen Validierungsstudie ab April 2025 an zwei Standorten des Universitätsklinikums Schleswig-Holsteins evaluiert. Hierbei werden u.a. die Prädiktionsgüte im Anwendungssetting und der zeitliche Unterschied zwischen Risikowarnung und dem Eintreten des medizinischen Ereignisses analysiert.

Ergebnisse:

In der retrospektiven Evaluation betrug die Klassifikationsleistung der trainierten Modelle für zehn medizinische Risiken zwischen 0,878 und 0,954. Die Warnungen von sieben entwickelten Risikoprädiktionsmodellen wurden im simulierten Anwendungsszenario als überwiegend angemessen und nachvollziehbar erklärt beurteilt. Erste Ergebnisse der Validierungsstudie werden im Beitrag präsentiert.

Diskussion:

Über den beschriebenen Weg konnten sieben ML-Prädiktionsmodelle für ein CDSS entwickelt werden, welches als Medizinprodukt der Klasse IIa zertifiziert ist. Ab April 2025 erfolgt die erste Post-Implementierungsstudie, weitere klinische Studien bezüglich der Veränderung von Prozessen und klinischen Endpunkten sollen folgen.

Referenzen:

1. Adams, R. et al. (2022) ‘Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis’, Nature Medicine, 28(7), pp. 1455–1460. Available at: <https://doi.org/10.1038/s41591-022-01894-0>.

Citation

Baehr, F., Ludolph, P., Dastur, M., Hugo, J., Schädler, D., Behn, C.-O., Wehkamp, K., & Ladenbauer, J. (2025, September 23). Entwicklung und Implementierung von KI-basierter klinischer Entscheidungsunterstützung in der stationären Versorgung. 2. DGDM Symposium für Digitale Medizin (2. DGDM), Hasso-Plattner-Institut, Potsdam. <https://doi.org/10.5281/zenodo.17058306>

Kennzahlenentwicklung zur ROI-Bewertung von KHZG-konformen Patientenportalen am Beispiel der No-Show-Rate in der Neurochirurgie

C. Hörning^{1*}, M.Holling²

*präsentierender Autor (bitte auch unterstreichen)

1 christian.hoerning@ukmuenster.de, Universitätsklinikum Münster, Klinik für Neurochirurgie

2 Universitätsklinikum Münster, Klinik für Neurochirurgie

DOI: 10.5281/zenodo.17058420

Einleitung:

Das Krankenhauszukunftsgesetz (KHZG) fördert mit dem Fördertatbestand 2 die Einführung digitaler Patientenportale zur Verbesserung der Effizienz und Patientenorientierung in der

Versorgung. Für eine belastbare Bewertung des Return on Investment (ROI) fehlen bislang standardisierte, quantitative Kennzahlen. Das Ziel dieser Arbeit ist die exemplarische Analyse der No-Show-Rate als operationalisierbarer Indikator zur ROI-Evaluierung.

Methoden:

In einer retrospektiven Sekundärdatenanalyse des Krankenhausinformationssystems (KIS) ORBIS wurden alle ambulanten Terminvereinbarungen der Klinik für Neurochirurgie am Universitätsklinikum Münster im Jahr 2024 (n = 8.475) untersucht. Neben der No-Show-Rate wurden die durchschnittliche Wartezeit auf einen Ersttermin (153 Tage) sowie administrative Folgekosten durch Neutermminierung (5 Minuten) analysiert. Simuliert wurde eine Reduktion der No-Show-Rate durch digitale Patienteninterventionen (z. B. automatisierte Erinnerungsfunktionen, Self-Service-Portale). Zur Abschätzung des Nutzens wurde eine hypothetische Reduktion der No-Show-Rate um 30 % bzw. 50 % angenommen. Daraus wurden Ressourcenpotenziale, Zeitgewinne und Auswirkungen auf die Wartezeiten modelliert.

Ergebnisse:

Die No-Show-Rate lag bei 5,12 %, entsprechend 433 versäumten Terminen. Durch digitale Maßnahmen wurde eine Reduktion auf 303 bzw. 217 Fälle (-30 % bis -50 %) modelliert. Dies entspricht einer Einsparung von jährlich rund 2.165 Minuten administrativen Aufwands (>36 Stunden) bei 5 Minuten je Neutermminierung. Gleichzeitig ergibt sich eine potenzielle Verkürzung der Wartezeit auf einen Ersttermin durch Wiederfreigabe geblockter Kapazitäten von 2,4 bis 3,9 Tage (-30 % bis -50 %). Hochgerechnet auf das gesamte Universitätsklinikum Münster mit ca. 500.000 ambulanten Kontakten jährlich ergibt sich ein potenzielles Einsparvolumen von über 208.000 Minuten administrativen Aufwands, was rund 3,6 Vollzeitäquivalenten entspricht (bei 38,5 Wochenstunden). Zudem könnten jährlich bis zu 12.800 Termine reaktiviert und in die Versorgung rückgeführt werden. Dies würde bei gleichbleibender Ressourcenausstattung eine relevante Reduktion der durchschnittlichen Wartezeiten von 5,6 bzw. 9,3 Tage (-30 % bis -50 %) auf einen Ersttermin klinikweit ermöglichen.

Diskussion:

Die No-Show-Rate eignet sich als praktikable und ökonomisch relevante Kennzahl zur ROI-Bewertung von Patientenportalen im KHZG-Kontext. Ihre Erhebung erlaubt eine differenzierte Einschätzung der Auswirkungen digitaler Lösungen auf Ressourcennutzung, Versorgungskapazität und Verwaltungsaufwand. Der Nutzen digitaler Terminmanagementsysteme reicht damit über Effizienzgewinne hinaus und hat strategische Bedeutung für den Ressourceneinsatz in Großkliniken.

Tastaturnutzung als digitaler Marker feinmotorischer Funktion bei Multipler Sklerose

P. Kirschner^{1*}, M. Korsen², S.G. Meuth², M. Pawlitzki², L. Masanneck²

*präsentierender Autor (bitte auch unterstreichen)

1 patricia.kirschner@med.uni-duesseldorf.de, Klinik für Neurologie, Universitätsklinikum Düsseldorf, Deutschland

2 Klinik für Neurologie, Universitätsklinikum Düsseldorf, Deutschland

DOI: 10.5281/zenodo.17058452

Einleitung: Multiple Sklerose (MS) ist eine chronisch-entzündliche Erkrankung des zentralen Nervensystems, die sowohl mit schubförmiger als auch schleichend-progredienter Krankheitsaktivität einhergehen kann. Neben Mobilitätseinschränkungen treten im Krankheitsverlauf oft kognitive und feinmotorische Defizite auf. Konventionelle Testverfahren wie der 9-Hole-Peg-Test (9HPT), der die Handmotorik erfasst, sind zwar in Studien etabliert, finden jedoch in der klinischen Routine selten kontinuierliche Anwendung. Digitale Methoden bieten das Potenzial, feinmotorische und kognitive Einschränkungen objektiv zu erfassen und somit schleichende Progredienz frühzeitig zu erkennen.

Methoden: Im Rahmen einer prospektiven Beobachtungsstudie (NCT06501950) wurde das Tastaturnutzungsverhalten von MS-PatientInnen über die Neurokeys-Tastatur-App auf iOS-/Android-Geräten analysiert. Anzahl und Dynamik der Tastenanschläge wurden erfasst, die Daten zunächst lokal gespeichert und anschließend in eine Cloud übertragen. Aus den digitalen Daten wurde ein geschätzter 9HPT-Score berechnet und zusätzlich ein manueller 9HPT durchgeführt. Die Korrelation beider Scores wurde mittels Pearson-Koeffizient bestimmt.

Ergebnisse: Das mediane Alter der 24 Teilnehmenden betrug 42,5 (22–58) Jahre. Der digitale 9HPT-Score an Tag 6 nach App-Installation korrelierte signifikant mit dem manuellen Score der dominanten Hand ($r=0,764$, $p<0,05$). An Tag 14 war die Korrelation geringer ($r=0,65$, $p<0,05$). Eine signifikante Korrelation bestand auch zwischen digitalem 9HPT-Score und Alter ($r=0,43$, $p<0,05$).

Diskussion: Die Auswertung der Neurokeys-Tastatur-Daten korreliert zufriedenstellend mit den Ergebnissen des im klinischen Alltag etablierten 9HPT. Gewöhnungseffekte bei der Tastaturnutzung beeinflussen jedoch das Ergebnis. Datenschutzbedenken der PatientInnen und die notwendige Gewöhnung an die neu installierte Tastatur limitieren die Nutzungsadhärenz. Passive, digitale Verfahren, wie die Neurokeys-Tastatur, ermöglichen eine kontinuierliche, alltagsnahe Erfassung von MS-Symptomen und können konventionelle Tests im klinischen Alltag und in Studien ergänzen.

Analyse der Schlafqualität mittels klinischer und digitaler Messmethoden bei Patienten mit chronischer inflammatorischer demyelinisierender Polyneuropathie

Jan Voth^{1*}, Charlotte von Gall², Noëmi Gmahl³, Noah M. Werner¹, Niklas Huntemann¹, Gerd Meyer zu Hörste³, Tobias Ruck⁴, Sven G. Meuth¹, Marc Pawlitzki¹, Lars Masanneck^{1,5}

*präsentierender Autor (bitte auch unterstreichen)

1 jan.voth@med.uni-duesseldorf.de, Klinik für Neurologie, Universitätsklinikum Düsseldorf, Düsseldorf, Deutschland

2 Institut für Anatomie II, Medizinische Fakultät, Heinrich-Heine-Universität, Düsseldorf, Deutschland

3 Klinik für Neurologie, Universitätsklinikum Münster, Münster, Deutschland

4 Klinik für Neurologie, BG Universitätsklinikum Bergmannsheil, Bochum, Deutschland

5 Hasso-Plattner-Institut, Universität Potsdam, Potsdam, Deutschland

DOI: 10.5281/zenodo.17080939

Einleitung: Patienten mit chronischer inflammatorischer demyelinisierender Polyneuropathie (CIDP) berichten neben motorischen und sensiblen Defiziten häufig über Fatigue, welche die Lebensqualität maßgeblich einschränkt, jedoch unzureichend erforscht ist. Wearables erfassen objektive Schlafparameter kontinuierlich und könnten vielversprechende Ergänzungen zu klinischen Scores darstellen.

Ziele: Die Electronic Monitoring of Disease Activity in Patients with Chronic Inflammatory Demyelinating Polyneuropathy (EMDA-CIDP) Studie untersucht neben körperlicher Aktivität die Schlafqualität von CIDP-Patienten mittels digitaler Messungen und vergleicht diese mit etablierten Scores und Fragebögen.

Methoden: In dieser prospektiven Studie wurden CIDP-Patienten der Universitätskliniken Düsseldorf und Münster 6 Monate beobachtet (NCT05723848).

Es erfolgten drei Visiten mit klinischen Scores wie Inflammatory Neuropathy Cause and Treatment Disability Score (INCAT) und Inflammatory Rasch-built Overall Disability Scale (I-RODS) sowie Fragebögen (Pittsburgh Sleep Quality Index, PSQI). Zusätzlich wurden Schlafparameter wie Schlaffeffizienz und Wachzeiten nach Schlafbeginn (WASO), sowie

Aktivitäts- und kardiale Parameter mithilfe einer Smartwatch (ScanWatch, Withings) aufgezeichnet.

Ergebnisse: Von 46 eingeschlossenen Studienpatienten wurden 40 (22,5% weiblich) in die Analyse einbezogen. Bei Studieneinschluss betrug das mediane Alter (Interquartilabstand) 66 Jahre (59,5; 72,3), der PSQI 6 (3; 8) und der INCAT Score 2 (1; 4). Die mediane nächtliche Adhärenz zur Smartwatch betrug 98,6% (95,8; 99,7). Es bestand eine moderate positive Korrelation zwischen PSQI und klinischer Symptomlast (Spearman: -0,43 I-RODS, 0,34 INCAT; jeweils $p < 0,05$) sowie zwischen PSQI und ausgewählten digitalen Schlafparametern (Spearman: -0,49 Schlafeffizienz, 0,36 WASO; jeweils $p < 0,05$).

Diskussion: Die EMDA-CIDP-Studie zeigt, dass CIDP-Patienten häufig eine eingeschränkte Schlafqualität aufweisen, welche sich digital erfassen lässt und mit der klinischen Symptomlast assoziiert ist. Wearables könnten zur objektiven Verlaufskontrolle beitragen.

Cost-effectiveness and -utility of a tele-neurologic intervention alongside a stepped wedge randomized controlled trial in a rural area within Germany (NeTKoH)

Paula Johanna Filser^{1,2*†}, Ricarda Sophia Schulz^{1†}, Kerstin Wainwright¹, Hans-Aloys Wischmann¹, Simone Kiel³, Tobias Kurth¹ & the NeTKoH Consortium

*presenter (also underline presenter please)

¹ paula.filser@charite.de, , Institute of Public Health, Charité – Universitätsmedizin Berlin, Germany

² Department of Healthcare Management, Technische Universität Berlin, Germany

³ Department of General Practice, Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany

† Paula Johanna Filser and Ricarda Sophia Schulz shared first authorship

DOI: 10.5281/zenodo.17081031

Introduction:

Neurological disorders contribute significantly and increasingly to the global burden of disease. Tele-neurology offers an opportunity to improve healthcare, particularly in rural and underserved regions. The health economic evaluation embedded within the NeTKoH trial assessed the economic impact of a tele-neurologic intervention in primary care in a rural region of Germany.

Methods:

Cost-effectiveness and -utility analyses were conducted to calculate incremental cost-effectiveness ratios (ICERs). For the cost-utility analysis the outcome of interest were quality-adjusted life years (QALYs) assessed via EQ-5D-5L. Cost-effectiveness analyses were conducted for the (1) proportion of neurological cases resolved in primary care, (2) number and (3) length of hospital stays, and (4) waiting time until consultation with a neurologist. Cost data were routinely collected by participating statutory health insurances and analyzed from their perspective.

Results:

372 participants were analyzed. The adjusted number of neurological cases resolved in primary care was -0.20 (95% CI [-0.31; -0.08]), while the incremental number of patients with at least one hospitalization was -0.05 [-0.11; 0.02], and the incremental effect on QALYs was 0.00 [-0.01; 0.01]. A significant reduction was observed in waiting time for a first neurologist contact (39.79

days, 95% CI [29.95; 49.45]), with an ICER of -5.79€ per day saved. Other cost-effectiveness outcomes showed no statistically significant effects, with slightly increased costs.

Discussion:

The analyses provide new insights and inform the discussion regarding the future use of tele-neurology in rural Germany and similar healthcare systems.

Digitale Navigator:innen zur Akzeptanz- und Kompetenzentwicklung mit Mental Health Apps: Ergebnisse einer Mixed-Method-Interventionsstudie (DigiNavi)

J. Schwarz^{1*}, D. Schubert¹, J. Schönbeck¹, K. Chen², J. Torous², J. Wolff³, J. Speck¹, E. Meier-Diedrich^{1#}, L. Uchtmann^{1#}

*präsentierender Autor; #geteilte Letztautor:innenschaft

1 julian.schwarz@mhb-fontane.de, Klinik für Psychiatrie und Psychotherapie, Zentrum für Seelische Gesundheit, Immanuel Klinik Rüdersdorf, Universitätsklinikum der Medizinischen Hochschule Brandenburg, Rüdersdorf, Deutschland

2 Division of Digital Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, USA

3 Peter L. Reichertz Institut für Medizinische Informatik (PLRI) der Medizinischen Hochschule Hannover, Braunschweig, Deutschland

DOI: 10.5281/zenodo.17081051

Einleitung: In Deutschland warten Patient:innen häufig mehrere Monate auf einen psychotherapeutischen Behandlungsplatz. Digitale Gesundheitsanwendungen (DiGAs) können zeitnah unterstützen, werden aber bislang nur überschaubar genutzt. Digitale Navigator:innen (DN) – speziell geschulte Gesundheitsfachkräfte – begleiten Patient:innen und Behandler:innen bei der Auswahl und Nutzung von DiGAs. Sie können so digitale Gesundheitskompetenz stärken, Engagement fördern und Fachkräfte entlasten. Die vom BMG geförderte DigiNavi-Studie (07/2024–09/2025) pilotiert den Einsatz von DN in der hausärztlichen und ambulanten psychiatrischen Versorgung. Ziel ist es, Erwartungen, Herausforderungen, Nutzungserfahrungen sowie Effekte auf die Versorgung zu untersuchen.

Methoden: Die Mixed-Methods-Machbarkeitsstudie wird an sechs Standorten (3 psychiatrische Institutsambulanzen, 3 Praxen für Allgemeinmedizin) durchgeführt. In einer Vorstudie werden Interviews und Fokusgruppen mit Patientinnen und Behandlerinnen qualitativ ausgewertet. Ein Manual der Harvard Medical School zum Training von DNs wird für den deutschsprachigen Raum adaptiert. N=6 Medizinische Fachangestellte werden zu DNs ausgebildet. In der Interventionsphase werden Patient:innen mit psychischen Störungen (ICD-10 F1-F6, F06.7) über drei Monate beim Einsatz von DiGAs begleitet. Die digitale Gesundheitskompetenz, Veränderungsbereitschaft, digitale/technische Kompetenz sowie Krankheitsschwere werden bei Patient:innen (n=48) und Behandler:innen (n=18) im Prä-Post-Design erhoben und vergleichend analysiert. Ergänzend werden qualitative Interviews mit allen Beteiligten durchgeführt.

Ergebnisse: Die Studie liefert Erkenntnisse zur Akzeptanz und Umsetzbarkeit hybrider Versorgung durch DN in multiprofessionellen Settings. Die Ergebnisse liegen bis Mitte 2025 vor und werden auf dem Symposium der DGDM vorgestellt.

Diskussion: Die Einführung von DN kann die Integration digitaler Angebote in die Versorgung fördern und sich positiv auf das Engagement, die Adhärenz und die Behandlungsergebnisse bei der Nutzung digitaler Anwendungen zur psychischen Gesundheit auswirken.

Interdisziplinäre Mustererkennung digitaler Phänotypen in der Multiplen Sklerose mittels maschinellen Lernens: ein Use Case aus Gang- und MRT-Daten.

Hernan Inojosa^{1*}, Wanqi Zhao², Judith Wenk¹, Isabel Voigt¹, Dirk Schriefer¹, Rocco Haase¹, Heidi Stölzer-Hutsch¹, Stefanie Fischer¹, Hagen Kitzler³, Hanes Schleiter², Maren Kählig², Uwe Aßmann², Katrin Trentzsch¹, Karsten Wendt², Tjalf Ziemssen¹

*präsentierender Autor

¹ hernan.inojosa@ukdd.de. Zentrum für Klinische Neurowissenschaften, Universitätsklinikum Carl Gustav Carus Dresden, Technische Universität Dresden, Fetscherstraße 74, Dresden 01307, Deutschland

² Professur für Softwaretechnologie, Technische Universität Dresden, 01062 Dresden, Deutschland

³ Institut für Diagnostische und Interventionelle Neuroradiologie, Universitätsklinikum Carl Gustav Carus Dresden, Technische Universität Dresden, Fetscherstraße 74, Dresden 01307, Deutschland

DOI: 10.5281/zenodo.17081096

Einleitung: Digitale Technologien eröffnen neue Möglichkeiten, neurologische Erkrankungen wie die Multiple Sklerose (MS) besser zu verstehen. Im interdisziplinär geführten Projekt „DigiPhenoMS“ an der TU Dresden wurde exemplarisch gezeigt, wie durch die enge Zusammenarbeit zwischen klinischen Expert*innen und Data Scientists digitale Phänotypen identifiziert werden können. Im Fokus stand der Aufbau eines datenschutzkonformen Analyse-Workflows zur Anwendung von Machine Learning mit dem Ziel, digitale Biomarker zu entwickeln. Wir berichten über die Rationale und erste Ergebnisse einer Clustering-Analyse mittels Gang- und MRT-Daten unserer MS-Kohorte.

Methoden: Zwischen 2018 und 2023 wurden von über 1.000 MS-Betroffenen quantitative Gang- und MRT-Daten erhoben. Definierte Parameter wurden pseudonymisiert und auf einem sicheren Server gespeichert. Kern des Projekts war ein interdisziplinärer kontinuierlicher Workflow: In gemeinsamen Workshops identifizierten Neurolog:innen, Bewegungsanalyst:innen, Radiolog:innen, Informatiker:innen und Data Scientists gemeinsam klinisch relevante Fragestellungen, wählten Variablen aus und stimmten die Datenaufbereitung ab. Die Analysepipeline umfasste eine Principal Component Analysis (PCA) zur Dimensionsreduktion sowie ein k-means-Clustering zur Mustererkennung. Zwischenergebnisse wurden fortlaufend im Team interdisziplinär reflektiert und kontextualisiert.

Ergebnisse: Es konnten vier unterscheidbare Mobilitäts-/MRT-Cluster identifiziert werden. Besonders auffällig war ein intermediäres Profil mit erhöhter Schrittfrequenz und kürzeren Schritten bei moderater Geschwindigkeit - ein Hinweis auf ein kompensatorisches Gangmuster.

Dieses Cluster zeigte strukturelle MRT-Veränderungen (u. a. geringere kortikale Volumina, höhere Läsionslast), trotz milder klinischer Beeinträchtigung.

Diskussion: Dieser Use Case zeigt das Potenzial digitaler Analyseverfahren zur Identifikation funktioneller MS-Phänotypen. Die Kombination aus klinischer Erfahrung und datenbasierter Perspektiven liefert neue Ansätze für personalisierte Versorgung und frühzeitige Intervention.

Safety and Security Implications of LxM-Enhanced and Agentic Systems in Healthcare

Oscar Freyer^{1*}, Jan Niklas Clusmann^{1,2}, Max Ostermann^{1*}

*presenter, oscar.freyer@tu-dresden.de, max.ostermann@tu-dresden.de

1 Else Kroener Fresenius Center for Digital Health, Technical University Dresden, Dresden, Germany

2 Department of Medicine III, University Hospital RWTH Aachen, Aachen, Germany

DOI: <https://doi.org/10.5281/zenodo.17137248>

Introduction: Foundation models (LxM) are anticipated to be integrated into healthcare systems, assisting with tasks ranging from clinical decision-making to patient monitoring and assistive care. Their capacity for human-like communication and adaptive responses to diverse inputs facilitates more natural and autonomous actions. However, as LxMs and agentic systems are deployed in safety-critical contexts such as digital health applications, ambient systems, and embodied agents like robots, they introduce various safety and security risks.

Methods: To identify these risks, we first defined tasks and potential applications of LxM-enhanced and agentic systems in healthcare. In the second step, we developed architectures and data flow maps for these applications, covering perception, processing, actuation, and network communication components. In a third step, we conducted a structured threat analysis using the STRIDE methodology and evaluated resulting clinically relevant risks based on established frameworks for medical risk management.

Results: In the analysis, we identified vulnerabilities arising from both the AI components (e.g., prompt injection, model poisoning) and system-level integration (e.g., spoofed sensors and insecure network interfaces). These vulnerabilities can lead to safety risks when LxMs directly influence physical actions or clinical recommendations. Agentic behavior amplifies these concerns by enabling autonomous actions with limited human oversight and increasing the likelihood of cascading errors.

Discussion: Ensuring the safe deployment of LxM-enhanced and agentic systems in healthcare requires anticipating not just individual failure points but also emergent risks arising from adaptive behavior and autonomous actions. Threat-informed design, robust validation, and adversarial testing will be key to mitigating harms while enabling responsible innovation.

Leveraging ChatGPT for Systematic Reviews - A feasibility study and framework proposal

I.Schnorr^{1,2*}, K. S. Appel^{1*}, J.J. Vehreschild^{1,2}, D. Maier^{1,3}

*presenter, contributed equally

1 i.schnorr@med.uni-frankfurt.de, Goethe University Frankfurt, Faculty of Medicine, Institute for Digital Medicine and Clinical Data Science, Germany

2 Goethe University Frankfurt, University Hospital, Center for Internal Medicine, Medical Department 2 (Hematology/Oncology and Infectious Diseases), Frankfurt, Germany

3 German Cancer Consortium (DKTK), partner site Frankfurt/Mainz, a partnership between DKFZ and University Medicine Frankfurt, Frankfurt am Main, Germany

DOI: <https://doi.org/10.5281/zenodo.17137336>

Introduction: Conducting a systematic review is intricate but indispensable to generate reliable evidence. While manual execution is resource-intensive [1], Large Language Models (LLMs) offer a scalable alternative. This study evaluates the feasibility and performance of an LLM workflow to automate systematic reviews in medical research.

Methods: Based on a published systematic review about clinical prognostic scores for COVID-19 [2], we developed and validated an LLM-supported technical framework structured according to PRISMA [3]. By employing detailed prompt engineering, targeting article screening and data extraction, we instructed GPT-4o to replicate key steps of the review process. Performance metrics were compared to the original human review results, which were considered as the ground truth.

Results: Preliminary results indicate that the LLM's title and abstract screening achieved high accuracy (91.58%) with a low False-Negative Rate (FNR; 6.81%) and a Cohen's Kappa of 0.75 (0.69-0.80), suggesting strong potential for initial article screening. The subsequent full-text screening demonstrated moderate overall accuracy (74.64%). The associated FNR of 27% indicates a significant risk of excluding relevant studies.

Discussion: The LLM-based approach exhibits strong potential to accelerate the article screening and to substantially reduce manual effort in successive steps. However, the performance drop observed during full-text screening reflects the contextual complexity of scientific articles. Our findings suggest that LLMs could serve best as a co-pilot in the systematic review process, augmenting researcher capabilities rather than offering a fully autonomous solution.

References

1. Affengruber L, van der Maten MM, Spiero I *et al.* *An exploration of available methods and tools to improve the efficiency of systematic review production: a scoping review.* *BMC Medical Research Methodology* 2024; 24:210. <https://doi.org/10.1186/s12874-024-02320-4>.
2. Appel KS, Geisler R, Maier D *et al.* A Systematic Review of Predictor Composition, Outcomes, Risk of Bias, and Validation of COVID-19 Prognostic Scores. *Clinical Infectious Diseases* 2024; 78 (4):889–899, <https://doi.org/10.1093/cid/ciad618>.
3. Page MJ, McKenzie JE, Bossuyt PM *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. <http://doi.org/10.1136/bmj.n71>.

Wie sehr stimmen Large Language Modells in ihren Empfehlungen zu Diagnostik und Therapie eines*r Patient*in mit den Empfehlungen der Leitlinie überein? Eine Betrachtung am Beispiel der S2K Leitlinie zum Hämolytisch-urämischem Syndrom

A. Gieser¹, S. Pudasaini², D. Müller³

1 adrian-markus.gieser@charite.de, Charité – Universitätsmedizin Berlin, Klinik für Pädiatrie mit Schwerpunkt Gastroenterologie, Nephrologie und Stoffwechselmedizin, Deutschland

2 Charité – Universitätsmedizin Berlin, Klinik für Pädiatrie mit Schwerpunkt Gastroenterologie, Nephrologie und Stoffwechselmedizin, Deutschland

3 Charité – Universitätsmedizin Berlin, Klinik für Pädiatrie mit Schwerpunkt Gastroenterologie, Nephrologie und Stoffwechselmedizin, Deutschland

DOI: 10.5281/zenodo.17137404

Einleitung: Die Fähigkeit von Large Language Modells (LLM) auf Anforderung beliebige Texte zu produzieren, Fragen wissenschaftlich zu beantworten oder längere Texte zusammenzufassen machen diese für einen Einsatz im Gesundheitswesen als Recherche Tool interessant. Diese Arbeit untersucht, inwiefern die Antworten zweier LLMs, exemplarisch ChatGPT™ und Perplexity™, in der medizinischen Beratung zur Diagnostik und Therapie eines Kindes mit hämolytisch-urämischem Syndrom (HUS) mit den Empfehlungen der entsprechenden AWMF-S2k-Leitlinie übereinstimmen.

Methoden: Im Rahmen der Arbeit wurden drei Fragensets entwickelt: eines basierend auf einer wörtlichen Übersetzung der Leitlinienempfehlungen, ein weiteres welches die Leitlinienempfehlungen freier übersetzt, sodass die Fragen dem klinischen Alltag eher entsprechen und ein drittes, das reale Patient*innenfälle abbildet. Die Antwortqualität der LLMs wurde dabei auf einer fünfstufigen Likert-Skala bewertet und mittels Kolmogorov-Smirnov- sowie Mann-Whitney-U-Tests statistisch untersucht.

Ergebnisse: Die Ergebnisse zeigen, dass ChatGPT™ in 70,59 % und Perplexity™ in 67,65 % der Fälle leitliniengerechte Antworten liefert, dabei konnte kein statistisch signifikanter Qualitätsunterschied zwischen den Modellen ($p = 0,728$) nachgewiesen werden. Qualitativ lassen sich teilweise Unterschiede in der Antwortstreuung feststellen. Weiterhin zeigte sich, dass Perplexity™ in manchen Antworten widersprüchliche Empfehlungen präsentiert.

Diskussion: Beide Modelle liefern in einer Mehrheit der Fälle leitlinienkonforme Antworten, erzielen, jedoch nicht durchgängig eine vollständige Übereinstimmung. Aufgrund der bestehenden

Fehlerquoten und der Notwendigkeit einer kontinuierlichen Quellenvalidierung, können LLMs derzeit nicht als alleiniges Recherchetool im klinischen Rahmen eingesetzt werden. Die Ergebnisse müssen im klinischen Kontext und in Kombination mit fundierter Fachliteratur überprüft werden. Die qualitativen Ergebnisse weisen auf die Bedeutung eines sorgfältigen Prompt Engineerings hin, um bessere und konsistentere Ergebnisse zu erzielen.

Towards Personalized Neuroscience: Evaluating Individual-Level Information in Neural Mass Models

Clemens Pellengahr*¹, Carlotta B.C. Barkhau*^{1,8}, Zheng Wang², Lukas Fisch¹, Ramona Leenings¹, Nils R. Winter¹, Jan Ernsting¹, Maximilian Konowski¹, Dominik Grotegerd¹, Susanne Meinert^{1,3}, Julia M. Hubbert¹, Judith Krieger¹, Tiana Borgers¹, Kira Flinkenflügel¹, Elisabeth J. Leehr^{1,4}, Frederike Stein⁵, Florian Thomas-Odenthal⁵, Paula Usemann⁵, Lea Teutenberg⁵, Igor Nenadić⁵, Benjamin Straube⁵, Nina Alexander⁵, Andreas Jansen⁵, Christian Porschen⁶, Tilo Kircher⁵, John D. Griffiths^{2,7}, Hamidreza Jamalabadi⁵, Udo Dannlowski¹, Tim Hahn¹

* Shared first authorship, presenter: clemens.pellengahr@uni-muenster.de

1 Institute for Translational Psychiatry, University of Münster, Germany

2 Department of Psychiatry and Institute of Medical Sciences, University of Toronto, Toronto, ON, Canada

3 Institute for Translational Neuroscience, University of Münster, Germany

4 Department of Clinical Psychology and Psychotherapy Georg-Elias-Müller-Institute of Psychology, Georg-August-University of Göttingen, Germany

5 Department of Psychiatry and Psychotherapy, University of Marburg, Germany

6 Department of Anaesthesiology, Intensive Care and Pain Medicine, University Hospital Münster, Germany

7 Krembil Centre for Neuroinformatics, Centre for Addiction and Mental Health, Toronto, Canada

8 Institute for Biomagnetism and Biosignalanalysis, University of Münster, Germany

DOI: <https://doi.org/10.5281/zenodo.17137553>

Introduction:

Macro-scale brain modeling with neural mass models (NMMs) simulates whole-brain dynamics, revealing oscillatory patterns and network interactions. Although connectome-based NMMs hold promise for personalized neuroscience, their ability to capture individual-specific neural signatures remains underexplored due to limited sample sizes and computational challenges. This study quantifies the individual-level information encoded in personalized NMMs using large-scale data.

Aims:

We simulated individual brain dynamics in a large cohort to assess whether NMM parameters retain unique information beyond empirical structural (SC) and functional connectivity (FC). The study evaluates model accuracy in subject identification and examines correlations between model parameters and demographic traits.

Methods:

Using a differentiable Reduced Wong-Wang (RWW) model on resting-state fMRI data from 1,444 subjects in the MACS study, three model variants were optimized:

- Globalized model (GM) with 4 parameters

- Localized model (LM) with 658 parameters
- Connectivity model (CM) with 23,875 parameters

Performance was measured by the explained variance between simulated and empirical FC and subject reidentification accuracy. ANOVA and OLS regression analyzed associations with age, gender, BMI, years of schooling, and IQ, while machine learning approaches evaluated predictive power.

Results:

Explained FC variance was 4.12% (GM), 19.16% (LM), and 56.57% (CM). Subject reidentification accuracy rose from <1% (GM) to 25.9% (LM) and 99.79% (CM). Despite high identification accuracy, correlations with demographic traits were weak (effect sizes: ($\eta^2_{\text{partial}} \leq 0.03$, standardized $\beta \leq 0.234$), indicating limited capacity for personalized trait prediction.

Conclusions:

While NMMs effectively simulate resting-state dynamics and individual FC patterns, their parameters inadequately predict cognitive or demographic traits. Future work should explore alternative NMM frameworks and longer fMRI sequences to improve personal neural signatures.

Migraine Research with Smartwatch Application (MIRA) - A

prospective migraine study with Wearables

R. Hagler^{1*}, N. Werner¹, M. Dauti¹, R. Jansen¹, S. Meuth¹, J.-I. Lee¹, M. Pawlitzki¹, L. Masanneck¹, T. Kölsche¹

*presenting author (ramona.hagler@med.uni-duesseldorf.de)

¹ – Department of Neurology, Medical Faculty and University Hospital Düsseldorf, Heinrich Heine University Düsseldorf, Germany

DOI: <https://doi.org/10.5281/zenodo.17137627>

Introduction:

Digital tools are increasingly used in clinical practice to continuously monitor health data, yet the application of wearable technology in primary headache research remains underexplored. The MIRA study is a prospective observational study that combines digital health tools with wearables to enhance our understanding of migraine pathophysiology and its impact on patients.

Methods:

Patients with episodic or chronic migraine are monitored over a six-month period with assessments at baseline (V1), 3 months (V2), and 6 months (V3). Participants are provided with a smartwatch (Withings® ScanWatch 2) to continuously monitor heart rate, body temperature variations, and activity levels while also documenting their migraine attacks, triggers, and symptoms via smartphone app. Clinical health questionnaires (e.g., WHOQOL, MIDAS) are used to assess quality of life and migraine-related disability.

Results:

A total of 55 patients have been enrolled (mean age: 43.1 ± 10.5 years; 96.15% female). Throughout the observation period, 90.9% of participants received botulinum toxin therapy every three months. Preliminary findings reveal high adherence to the smartwatch (median wear time: 96.7%). Single case analyses show that e.g. alterations in sleep patterns, as captured by the smartwatch were subsequently followed by migraine attacks, possibly identifying migraine triggers and early warning signs. A systematic evaluation of these observations is currently underway and will be integrated into the final study results.

Discussion:

The MIRA study demonstrates the potential of integrating wearable devices with digital health tools to identify migraine triggers through continuous monitoring. This approach provides a longitudinal perspective on migraine dynamics supporting the development of individualized preventive strategies.

Bewertung der Akzeptanz von künstlicher Intelligenz in der Kinder- und Jugendmedizin – eine qualitative Fokusgruppenstudie mit Kindern und Jugendlichen

L. Reinhart^{1*}, J.-L. Kerth¹, A. C. Bischops¹, T. Meißner¹

¹ Lisa.reinhart@med.uni-duesseldorf.de, Klinik für Allgemeine Pädiatrie, Neonatologie und Kinderkardiologie, Universitätsklinikum Düsseldorf

DOI: <https://doi.org/10.5281/zenodo.17137650>

Einleitung: Mit zunehmender Digitalisierung findet Künstliche Intelligenz (KI) immer mehr Einsatz in der Kinder- und Jugendmedizin. Der langfristige Erfolg von KI hängt jedoch von der Akzeptanz der beteiligten Akteure ab, zu der bislang nur wenige Studien existieren. Diese qualitative Studie untersucht daher die Akzeptanz von KI durch Kinder und Jugendliche.

Methoden: Von März bis September 2024 wurden Kinder und Jugendliche (8-17 Jahre) am Zentrum für Kinder- und Jugendmedizin des Universitätsklinikums Düsseldorf durch E-Mails, Aushänge und persönliche Ansprache rekrutiert. Anhand eines Interviewleitfadens führten jeweils zwei Mitglieder des Studienteams Fokusgruppeninterviews durch, die zwischen 30-90 Minuten dauerten. Mittels allgemeiner Fragen zu KI und konkreter Szenarien wurden unter anderem Datenerhebung, Datenschutz, Nutzen, Ängste, Verantwortung, Transparenz, Umgang mit Fehldiagnosen, Recht auf Nichtwissen und der Einfluss auf das Arzt-Patienten-Verhältnis diskutiert. Die audioaufgezeichneten und transkribierten Interviews wurden durch zwei Mitglieder des Studienteams ausgewertet.

Ergebnisse: Es wurden 20 Fokusgruppeninterviews mit 60 Kindern und Jugendlichen durchgeführt. Vier Interviews waren mit Kindern (8-11 Jahre, n=11), sechs mit Jugendlichen (12-14 Jahre, n=23) und zehn mit Jugendlichen (15-17 Jahre, n=26). Die Früherkennung von Krankheiten, Betreuung chronischer Erkrankungen, Vernetzung mit Fachpersonal, Fehlervermeidung und niedrigschwellige Abklärung von Symptomen wurden als Vorteile gesehen. Bedenken gab es hinsichtlich des Datenschutzes, der Verantwortung für KI-Empfehlungen und fehlerhaften Befunden.

Diskussion: Es zeigt sich eine positive Einstellung gegenüber KI in der Kinder- und Jugendmedizin. Insbesondere in der Diagnostik, Prävention und Optimierung medizinischer Versorgung sehen Kinder und Jugendliche viel Potenzial. Datenschutzbedenken und mögliche Haftungsfragen stellen jedoch Hindernisse für die Akzeptanz da.

Commercialization of DiGA in Germany: Success Factors, Barriers, and Strategic Implications

L. Schramm*¹, C. Greis², F. C.-C. Carbon¹

*presenter (also underline presenter please)

¹ lukas.schramm@uni-bamberg.de, Department of General Psychology and Methodology, University of Bamberg, Bamberg, Bavaria, Germany

² University Hospital Zurich, Department of Dermatology, Zurich, Switzerland

DOI: <https://doi.org/10.5281/zenodo.17137664>

Introduction:

Digital health applications (DiGA) in Germany are evidence-based medical devices integrated into routine care. Despite or due to regulatory support, the number of DiGA manufacturers has barely grown and multiple manufacturers have gone bankrupt. On the other hand, the top 10 DiGAs account for 71% of the market volume. This study aims to identify key factors influencing DiGA success and to assess whether manufacturers with high prescription volumes differ strategically from those with lower volumes.

Methods:

A mixed-methods, Delphi-like design combined semi-structured expert interviews ($n = 32$) with a quantitative survey ($n = 142$) among DiGA manufacturers, healthcare professionals, and industry experts. Qualitative results informed the development of a structured questionnaire. Participants rated success factors, barriers, and strategies on a 10-point scale. Responses were compared between high- and low-prescription volumes manufacturers.

Results:

The most critical success factors were robust clinical evidence, strong physician networks, and awareness at the point of care. Key barriers included regulatory complexity and unawareness from physicians. Manufacturers with low-prescription volumes rated barriers and strategies significantly higher and more uniformly than high-prescription volumes manufacturers, who showed greater scoring variability.

Discussion:

This study delivers the first structured assessment of commercial success factors for DiGA. It emphasizes the importance of clinical evidence and access to the right physicians as primary success factors and regulatory burdens as key barriers. Strategic differentiation based on therapeutic area is essential based on the comprehensive analysis between high- and low-prescription DiGA manufacturer. These insights might support manufacturers and policymakers in improving DiGA implementation.

Digital Influence: Kategorisierung deutscher Medfluencer*innen auf Instagram

P. Tomasini^{1*}

*präsentierende Autorin (bitte auch unterstreichen)

¹ paula.tomasini@uni-jena.de, Universitätsklinikum Jena, Europäische Hochschule für Innovation und Perspektive (EHIP), Deutschland

DOI: <https://doi.org/10.5281/zenodo.17140969>

Einleitung:

Medfluencer*innen – Influencer*innen, die Inhalte zu gesundheitsbezogenen Themen veröffentlichen – gewinnen zunehmend an Bedeutung für die öffentliche Gesundheitskommunikation. Ihre Inhalte bewegen sich im Spannungsfeld zwischen Wissensvermittlung und wirtschaftlichen Interessen. Die aktuelle Berichterstattung befasst sich vorwiegend mit stark emotional aufgeladener, teils berechtigter Kritik an medizinisch aktiven Influencer*innen. Doch betrifft dies den gesamten Bereich? Welche Arten von Medfluencer*innen gibt es überhaupt? Ziel dieser Arbeit war es, eine systematische Kategorisierung deutscher Medfluencer*innen vorzunehmen, um eine Basis für einen konstruktiven Diskurs über positive und negative Einflüsse auf die Gesundheitskommunikation und -versorgung zu ermöglichen.

Methoden:

Untersucht wurden 8763 Beiträge von 100 deutschen Medfluencer*innen auf Instagram. Erfasst wurden unter anderem der Anteil wissensvermittelnder Beiträge, der Anteil werblicher Inhalte sowie Reichweitenkennzahlen. Eine Gruppierung erfolgte anhand inhaltlicher und statistischer Auswertung.

Ergebnisse:

Die Analyse zeigte teils signifikante Unterschiede zwischen den Gruppen: Medfluencer*innen mit Approbation publizierten einen höheren Anteil wissensvermittelnder Inhalte mit durchschnittlich größerer Fachlicher Tiefe. Influencerinnen ohne Approbation wiesen dagegen eine stärkere Verwendung von Affiliate-Links und Rabattcodes auf. Kein Unterschied fand sich beim Anteil klassischer Werbung. Insgesamt konnten 5 Medfluencer*innen-Typen abgeleitet werden, welche unterschiedliche Rollen im Gesundheitsdiskurs einnehmen.

Diskussion:

Die Ergebnisse bieten eine Grundlage für die systematische Einordnung von Medfluencer*innen. Sie eröffnen neuen Diskussionsspielraum, um dem größer werdenden Einfluss der digitalen

Gesundheitskommunikation gerecht zu werden. Darüber hinaus ergeben sich Erfolgsfaktoren für sinnvolle Kommunikation medizinischen Wissens, welche mit Standards der evidenzbasierten Medizin übereinstimmen und die Patient*innensicherheit vergrößert.

Who to Trust: the Patient or the Data? A 3-Month Telemonitoring Study in Youths with Post-COVID Syndrome

J. Greenfield^{1*}, Z. S. Oftring^{1,2}, J. Schmidt³, M. Hägele⁴, A. Farzaneh³, E. Hamelmann³, U. Behrends⁴, S. Kuhn¹

*presenter

1 julia.greenfield@uni-marburg.de, Institute for Digital Medicine, Philipps University Marburg and University Hospital Marburg, Marburg, Germany

2 Department of Paediatrics, University Hospital Marburg, Marburg, Germany

3 University Clinic OWL, Evangelisches Klinikum Bethel, Department of Paediatrics, Bielefeld, Germany

4 Technical University of Munich, University Hospital rechts der Isar, Clinic and Outpatient department for Paediatrics, MRI Chronic Fatigue Centrum (MCFC) for Adolescents, Munich, Germany

DOI: <https://doi.org/10.5281/zenodo.17141004>

Introduction: The “coverChild Telemonitoring” study recruited 30 paediatric patients (aged 12-17 years) with post Covid syndrome to undergo a 3-month app- and sensor based telemonitoring programme. This work presents the results of patient-reported and sensor-based data, focusing on well-being and spirometry.

Methods: The Fatigue Severity Scale (FSS) and Bell-Score were completed at 4-week intervals while at-home spirometry (FEV1 and PEF) was performed daily. Presence of respiratory symptoms was assessed at study start using the Munich Berlin Symptom Questionnaire (MBSQ). Data analysis compared subjective well-being and objective lung function data between study start and end.

Results: Respiratory symptoms were present in 20% (6/30) of patients. Subjective well-being improved significantly over the three months (Bell-Score increase: 15pts; FSS decrease: 0.25), while spirometry data showed significant decreases (FEV1 decrease: 0.25l; PEF decrease: 82.2ml) (Figures 1 and 2).

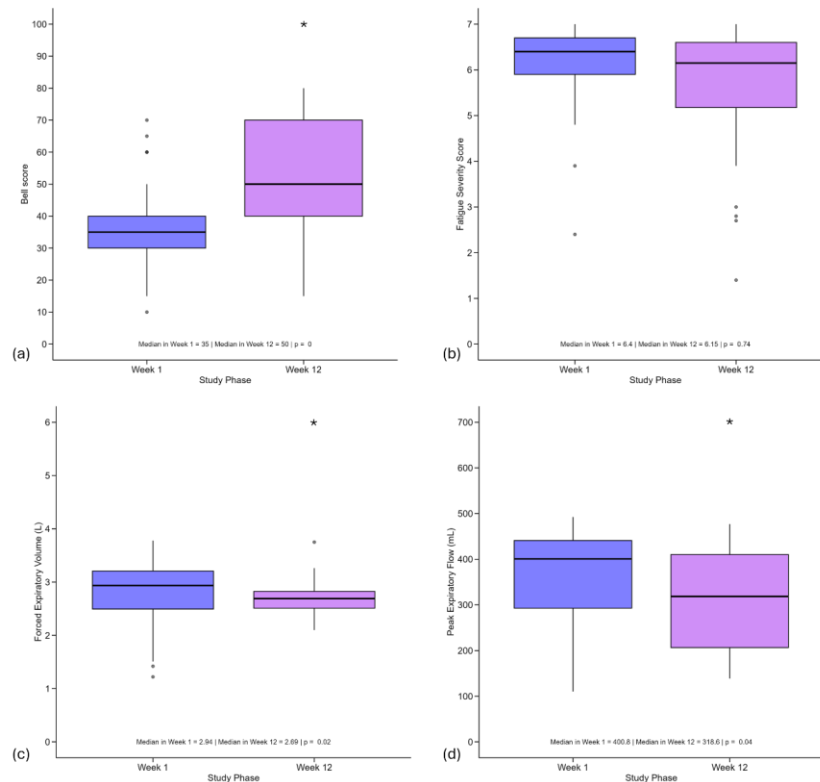


Figure 1. Progression of (a) Bell-Score and (b) FSS against (c) FEV1 and (d) PEF in all patients. Significant changes are marked with *.

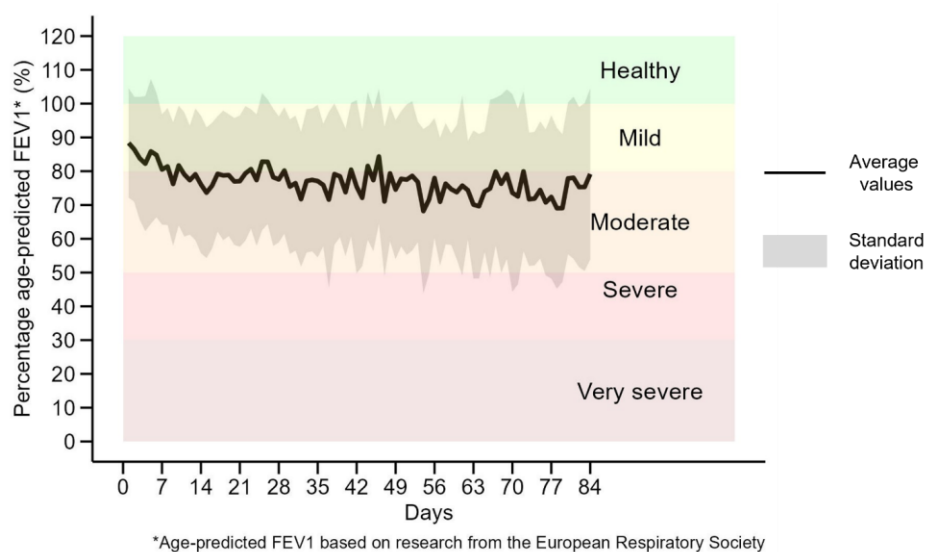


Figure 2. Percentage of age-predicted FEV1 values for all patients over the 3-month study period.

Discussion: The reported improvement in daily activity and decrease in fatigue was contradicted by spirometry data. Thresholds established by the European Respiratory Society provide

indications of healthy lung function, rendering the measured values of concern if taken at face-value. At-home spirometry has previously been criticised for lack of validity, and this work underscores the importance of validating sensor-based data and ensuring consistent and correct device usage. Future telemonitoring protocols may benefit from real-time data verification and repeated patient-education to resolve discrepancies.

Large language models for dermatological image interpretation – a comparative study

Lasse Cirkel^{*,1,2,3}, Fabian Lechner^{2,3}, Lukas Alexander Henk^{3,4}, Martin Krusche⁵, Martin C. Hirsch², Michael Hertl⁴, Sebastian Kuhn³, Johannes Knitza^{3,6}

*presenter

1 lcirkel@students.uni-mainz.de

2 Institute of Artificial Intelligence, University Hospital Gießen-Marburg, Philipps University, Marburg, Germany.

3 Institute for Digital Medicine, University Hospital Gießen-Marburg, Philipps University, Marburg, Germany.

4 Department of Dermatology and Allergology, University Hospital Gießen-Marburg, Philipps University, Marburg, Germany.

5 Division of Rheumatology and Systemic Inflammatory Diseases, III. Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

6 Université Grenoble Alpes, AGEIS, Grenoble, France.

DOI: <https://doi.org/10.5281/zenodo.17141082>

Introduction: Interpreting skin findings can be challenging for both laypersons and clinicians. Large Language Models (LLMs) offer accessible decision support, yet their diagnostic capabilities for dermatological images remain underexplored. This study evaluated the diagnostic performance of LLMs based on image interpretation of common dermatological diseases.

Methods: 500 dermatological images, encompassing four prevalent skin conditions (psoriasis, vitiligo, erysipelas and rosacea), were used to compare seven multimodal LLMs (GPT-4o, GPT-4o mini, Gemini 1.5 Pro, Gemini 1.5 Flash, Claude 3.5 Sonnet, Llama3.2 90B and 11B). A standardized prompt was used to generate one top diagnosis.

Results: The highest overall accuracy was achieved by GPT-4o (67.8%), followed by GPT-4o mini (63.8%) and Llama3.2 11B (61.4%). Accuracy varied considerably across conditions, with psoriasis with the highest mean LLM accuracy of 59.2\% and erysipelas demonstrating the lowest accuracy (33.4%). 11.0% of all images were misdiagnosed by all LLMs, whereas 11.6% were correctly diagnosed by all models. Correct diagnoses by all LLMs were linked to clear, disease-specific features, such as sharply demarcated erythematous plaques in psoriasis. Llama3.2 90B was the only LLM to decline diagnosing images, particularly those involving intimate areas of the body.

Discussion: LLM performance varied significantly, emphasizing the need for cautious usage. Notably, a free, locally hostable model correctly identified the top diagnosis for approximately

two-thirds of all images, demonstrating the potential for safer, locally deployed LLMs. Advancements in model accuracy and the integration of clinical metadata could further enhance accessible and reliable clinical decision support systems.

Translating Neurologist Notes into Disability Scores: Developing a Large Language Modell for Multiple Sclerosis Research and Care

Hernan Inojosa^{1*}, Ahmadreza Ramezanzadeh^{1,2}, Isabella C. Wiest², Dyke Ferber², Jakob Nikolas Kather², Tjalf Ziemssen^{1,3}

*presenter

1 hernan.inojosa@ukdd.de, Center of Clinical Neuroscience, Department of Neurology, Medical Faculty and University Hospital Carl Gustav Carus, TUD Dresden University of Technology, Germany

2 Else Kerner Stiftung for Digital Health, Medical Faculty and University Hospital Carl Gustav Carus, TUD Dresden University of Technology, Germany

DOI: <https://doi.org/10.5281/zenodo.17141150>

Introduction: Clinical documentation often relies on unstructured free text, particularly in routine practice and among non-specialists, limiting the systematic use of structured measures, such as the Expanded Disability Status Scale, the gold standard for quantifying disability in multiple sclerosis. However, scoring is time-consuming and requires specific training. This project aims to develop an artificial intelligence-based tool that converts narrative neurological descriptions into structured EDSS scores using large language models (LLMs).

Methods: We developed a locally deployed, fine-tuned LLM based on the Neurostatus-EDSS scoring framework. The model was trained to interpret free-text neurological examinations and output structured scores for EDSS-relevant functional systems (e.g. visual, pyramidal, cerebral) and EDSS, using a custom JSON schema. Training used anonymized clinical texts reflecting real-world documentation. The system includes a clinician-facing input module and a scoring engine leveraging the LLM.

Results: The tool generates plausible EDSS scores from free-text inputs, showing good agreement with expert ratings in domains with clear criteria, such as visual function and ambulation. Variability is greater in systems like pyramidal and cerebellar systems, where scoring and interpretation is more nuanced. The model also flags missing or ambiguous information and prompts users for clarification.

Discussion: Our project illustrates the potential of LLM-based EDSS scoring from narrative exams. While neurological examination remains clinician-drive, with advantages in flexibility but variability in documentation, this approach may support more consistent and structured assessments without replacing expert judgment. The underlying pipeline is

transferable to other clinical domains where free-text documentation limits structured data use.

Bedarfe und Wünsche von Hausärzten zu Informationen über digitale Gesundheitsanwendungen im Bereich Psyche

Fatma Sahan M.Sc.^{1*}, Charlotte Wagenaar¹, Karin Panitz M.Sc.¹, Prof. Dr. Nadja Kairies-Schwarz², Dr. Markus Vomhof², Maximilian Zinn M.Sc.², PD Dr. Jennifer Apolinário-Hagen¹

*präsentierender Autor (bitte auch unterstreichen)

¹ Fatma.Sahan@hhu.de, Institut für Arbeits-, Sozial- und Umweltmedizin, Heinrich-Heine-Universität Düsseldorf und Universitätsklinikum Düsseldorf, Düsseldorf, Deutschland

² Institut für Versorgungsforschung und Gesundheitsökonomie, Heinrich-Heine-Universität Düsseldorf und Universitätsklinikum Düsseldorf, Deutsches Diabetes Zentrum, Düsseldorf, Deutschland

DOI: <https://doi.org/10.5281/zenodo.17141177>

Einleitung: Hausärzte stellen häufig die erste Anlaufstelle für Menschen mit psychischen Beschwerden dar. Digitale Gesundheitsanwendungen (DiGA) bieten ein vielversprechendes Potenzial zur Unterstützung der hausärztlichen Versorgung, werden jedoch bislang zurückhaltend verschrieben. Die aktuelle Studienlage zeigt eine zunehmende Offenheit gegenüber der Verordnung, wobei Informationsdefizite als zentrale Hürde für die Integration in die Praxis gelten. Ziel dieser Studie ist es, Informationsbedarfe und -wünsche von Hausärzten hinsichtlich DiGA im Bereich Psyche zu untersuchen.

Methoden: Über semistrukturierte qualitative Interviews wurden Hausärzte im Rahmen einer DFG-geförderten Mixed-Methods-Studie deutschlandweit befragt. Transkribierte Interviews wurden mithilfe eines deduktiv-induktiven Kategoriensystems unter Nutzung von MAXQDA inhaltsanalytisch ausgewertet. Die Analyse konzentrierte sich auf die Kategorien Informationsquellen, Informationsinhalte, Informationsmedien sowie regulatorische Unklarheiten.

Ergebnisse: Insgesamt wurden die Daten von 13 Hausärzten (männlich: 77 %) mit einem Durchschnittsalter von 45 Jahren ($M=44.92$, $SD=14.63$; Median=54) analysiert. Bevorzugte Informationsquellen umfassten u. a. Ärztekammern, Berufsverbände und Krankenkassen. Informationsinhalte betrafen die Evidenzlage, Funktionalitäten und (Kontra-) Indikationen. Als Informationswege wurden z. B. Fachzeitschriften, Fachportale, Praxissoftware, Testzugänge sowie CME-Fortbildungen (digital oder analog) genannt. Informationen zu DiGA sollten im klinischen Alltag schnell und einfach abrufbar sein. Zu den berichteten regulatorischen Unklarheiten zählten die Sorge vor Regressforderungen und Fragen zur Budgetierung in der hausärztlichen Praxis.

Diskussion: Hausärzte wünschen sich leicht zugängliche, neutrale Informationen zu Evidenzbasis, Indikationen und Budgetierung von DiGA. Praxisnahe Entscheidungshilfen, integriert in

bestehende Praxissoftware, könnten die Informationsbeschaffung erleichtern. Vertrauenswürdige (industriunabhängige) Informationsquellen wie Ärztekammern können eine zentrale Rolle bei der Informationsdiffusion einnehmen. Unsicherheiten zur Budgetierung und Abrechnung verdeutlichen den Bedarf an administrativer Unterstützung der hausärztlichen Praxen.

Regulatorische Rahmenbedingungen und Fallbeispiele für multizentrische, dezentrale komplett digitale Studien mit Medizinprodukten

A. Schraplau¹, M. Kreher¹, L. Keller², S. Tiede¹, K. Piepelow¹, Caroline Schmidt-Lucke^{1*}

*präsentierender Autor (bitte auch unterstreichen)

1 Caroline.Schmidt-Lucke@mediacc.org, MEDIACC GmbH, Deutschland

2 Smoke Free 23 GmbH, Deutschland

DOI: <https://doi.org/10.5281/zenodo.17141221>

Einleitung:

Die digitale Umsetzung klinischer Prüfungen mit Medizinprodukten durch Einsatz dezentraler Elemente eröffnet disruptive Chancen zur Effizienzsteigerung, flächendeckenden Rekrutierung und Patient:innenzentrierung. Voraussetzung ist ein klarer rechtliche Rahmen, um Sicherheit, Datenschutz, Compliance und Retention sicherzustellen. Es wurde untersucht, unter welchen regulatorischen Rahmenbedingungen komplett digitale, formal multizentrische klinische Prüfungen bereits umsetzbar sind.

Methoden:

Zunächst wurden aktuelle Regulatorik und Leitlinien zu dezentralen, digitalen klinischen Prüfungen (DCT) mit Medizinprodukten in Europa evaluiert. Mittels systematischer Recherche wurden Studienprotokolle von Studien mit Medizinprodukten, insbesondere Digitale Gesundheitsanwendungen (DiGA) analysiert. Abschließend wurden praktische Ansätze zur Implementierung digitaler Lösungen entwickelt und in einer komplett digitalen multizentrischen, randomisiert-kontrollierten Studie veranschaulicht.

Ergebnisse:

Grobe Orientierung für den Einsatz dezentraler Elemente in klinischen Prüfungen mit Medizinprodukten finden sich in ISO14155, Dokumenten der MDCG und in Leitlinien und Gesetzen für Arzneimittelstudien sowie im Bundesmantelvertrag für Ärzte. Die Analyse von 62 publizierten bzw. registrierten klinischen Studien mit Medizinprodukten (inkl. 56 DiGA) ergab, dass einzelne digitale Elemente hauptsächlich für Rekrutierung, Datenerfassung und Einwilligung verwendet wurden. Risikoadaptiert wurde unter Berücksichtigung von Minimalanforderungen die komplett digital durchgeführte RAUCHFREI-Studie entwickelt. 14.500 Interessierte wurden automatisiert gescrout, 1466 Personen aus 16 Bundesländern digital eingeschlossen und

verblieben mit einer Retentionsrate von 84% nach 6 Monaten (Vergleich zum traditionellen Studienansatz 72%) in der Studie.

Diskussion:

Die überwiegend hybride Nutzung digitaler Elemente in analysierten Studien (v. a. Rekrutierung, eConsent) zeigt, dass regulatorische Unsicherheiten eine vollständige Dezentralisierung behindern. Die Ergebnisse unterstreichen das transformative Potenzial von DCT. Die erhöhte Retentionsrate in einer kompletten DCT deutet darauf hin, dass patientenzentrierte Ansätze und automatisierte Prozesse die Teilnehmendenmotivation und Effizienz steigern.

Screening von Publikationen mittels künstlicher Intelligenz: wie valide sind KI-extrahierte Studiendaten?

S. Eisenhofer^{1*}, S. Barke-Burjanko¹

*präsentierende Autorin: eisenhofer@gdekk.de

¹ Klinische Pharmazie, AMTS und E-Health, EKK plus GmbH, Deutschland

DOI:

<https://doi.org/10.5281/zenodo.17141246>

Einleitung:

Tools, die künstliche Intelligenz (KI) nutzen, bieten diverse Möglichkeiten Medizin und Pharmazie im klinischen Alltag zu unterstützen. Ihre Nutzung unterliegt jedoch Hürden und Grenzen, welche wissenschaftlich zu untersuchen sind. Die vorliegende Studie beschäftigt sich mit der Validität von Studiendaten, die durch ein KI-Tool aus wissenschaftlichen Publikationen extrahiert wurden.

Methoden:

Wissenschaftliche Publikationen zum Thema Dosierung von niedermolekularen Heparinen und Fondaparinux bei adipösen Patient*innen wurden durch eine Klinische Apothekerin zusammengefasst. Anschließend wurde ein umfassender Prompt für den PDF-Chat eines KI-Tools (scienceOS, scienceOS UG, Dresden) generiert, der die KI anweist folgende Parameter strukturiert aus den Publikationen zu extrahieren: Titel, Erstautor*in, Jahr, Studienart, Wirkstoff(e), Indikation, Studienpopulation, BMI bzw. Körpergewicht, Dosierung, Prozentsatz der erreichten/unterschrittenen/überschrittenen Anti-Xa-Target-Level, Anzahl und Schweregrad von Blutungen und Anzahl thrombovenöser Ereignisse (VTE). Die KI-extrahierten Daten wurden mit jenen der Klinischen Apothekerin verglichen und die Übereinstimmungsrate ermittelt.

Ergebnisse:

Es wurden je 15 Parameter aus 37 Publikationen verglichen und eine mediane Übereinstimmungsrate von 89 % (Q25/75 81 %/92 %; Min/Max 76 %/97 %) erreicht. Als besonders zuverlässig erwiesen sich die Parameter „Jahr“ und „Studienart“ (>90 %). Die höchste Übereinstimmungsrate (97 %) erzielten die Parameter „Anzahl VTE“ und „Indikation“. „Indikation“ wurde durch den Prompt in „VTE Therapie bzw. Prophylaxe“ vordefiniert. Schwächen wiesen die Parameter bezüglich „Erstautor*in“ (84%), Anti-Xa-Level (81-89 %), „Schweregrad von Blutungen“ (76 %) und „Dosierung“ (78 %) auf.

Diskussion:

Die Ergebnisse zeigen auf, dass KI-Tools eine wertvolle Unterstützung bei der Extraktion von Studiendaten aus wissenschaftlichen Publikationen darstellen können. Eine fachliche Überprüfung durch Expert*innen der Gesundheitsberufe, z.B. einer pharmazeutischen Arzneimittelinformation, ist dennoch unerlässlich.

Augmented Emotional Reality – Multimodale Emotionsübermittlung in der Telemedizin

F. Kegel^{1*}, L. M. Schlieper², M. Stein³, T. Kesztyüs², L. M. Kolbe¹, K. Schnell²

*präsentierender Autor

1 felix.kegel@uni-goettingen.de, Georg-August-Universität Göttingen, Deutschland

2 Universitätsmedizin Göttingen, Deutschland

3 Universität Heidelberg, Deutschland

DOI: <https://doi.org/10.5281/zenodo.17141274>

Einleitung: Die Telemedizin bietet zwar einen grundsätzlich adäquaten Ersatz zur Face-to-Face-Therapie in räumlicher Präsenz, verstärkt jedoch die typischen Probleme therapeutischer Kommunikation. Die reduzierte Bandbreite sensorischer Information erschwert die therapeutische Beziehungsgestaltung, da patientenseitige Zeichen für Kommunikationsbrüche schwerer erkennbar sind. Das BMBF-geförderte MITMed-Projekt entwickelt daher ein *Augmented Emotional Reality (AER)-System*: Patient:innen können ihren Therapeut:innen niederschwellig Veränderungen ihres emotionalen Zustands über einen zusätzlichen sensorischen Kanal mitteilen; ein am Unterarm der Therapeut:innen getragenes Wearable übermittelt die erfassten Emotionen als intuitiv wahrnehmbare taktile Reize. Ziel der AER ist die Stärkung der Arbeitsbeziehung.

Methoden: Das AER-System wird zunächst in einem prototypischen telemedizinischen Nutzungskontext in einer Grundlagenstudie getestet. Dabei werden im Rahmen der therapeutischen Ausbildung die Kontexte des Anamnesegesprächs zwischen Präsenz, Videogespräch und Videogespräch mit AER-Wearable variiert und verglichen. Eine weitere Studie untersucht als klinische Pilotanwendung den Nutzen der AER durch Vergleich der Videotherapie depressiver Störungen mit und ohne AER-Wearable.

Ergebnisse: Über die etablierten Dimensionen Valenz und Arousal ist eine intuitive parametrische Encodierung emotionaler Zustände auch taktil möglich, die wenig mit den audiovisuellen Videokontakt interferiert. Die Nutzererfahrungen der Grundlagenstudie deuten an, dass die Wearable-basierte AER die telemedizinische Kommunikation hinsichtlich der Entstehung und Reparatur von Kommunikationsbrüchen unterstützt. Zudem werden die Nutzererfahrungen von Patient:innen und klinischem Personal in der aktuell noch laufenden klinischen Pilotstudie vorgestellt.

Diskussion: Die AER über taktile Reize ermöglicht die intuitive Kommunikation über emotionale Zustände im telemedizinischen Kontext. Die teilweise noch laufenden Studien liefern erste Belege für den Nutzen neuer Kommunikationskanäle in klinischen Interaktionen. Es ergeben sich weitere vielversprechende Anwendungsfelder der AER im medizinischen Kontext.



Medical Billing Data: A hidden treasure for predicting Post-Endoscopic Mucosal Resection Adverse Events?

Moritz Ludwar^{1*}, Marcus Buchwald², Stefan Wittlinger¹, Sebastian Belle³

*presenter

1 moritz.ludwar@stud.uni-heidelberg.de, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

2 Mannheim Institute for Intelligent Systems in Medicine, Mannheim, Germany

3 Department of Medicine II, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

DOI: <https://doi.org/10.5281/zenodo.17141322>

Introduction: Medical billing data contains detailed and standardized information on clinical procedures, presenting an opportunity for secondary use in predictive analytics. This study aims to evaluate the potential of machine learning algorithms in predicting adverse events following endoscopic mucosal resection (EMR) using medical billing data.

Methods: The dataset comprised 3,468 EMR records from June 2008 to June 2022. Feature selection was performed based on data completeness and clustering techniques. Several classification algorithms, including Support Vector Machine, XGBoost, and LightGBM, were employed to predict adverse event risk. Model interpretability was assessed using SHAP analysis.

Results: Initial model accuracy ranged from 40% to 55%, indicating modest predictive performance.

Discussion: Current models do not yet achieve sufficient accuracy for clinical application. However, enhancing feature engineering and incorporating information extracted from free-text reports using Natural Language Processing could significantly improve model performance. Further research is needed to unlock the whole predictive potential of medical billing data.

Acceptance of and Engagement with VitaLo

J. Stein^{1,2*}, L. Roth²

*presenter

1 julian.stein@ukdd.de, Department for Prevention and Care of Diabetes, Department of Medicine III, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

2 Department for Prevention and Care of Diabetes, Department of Medicine III, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

DOI: <https://doi.org/10.5281/zenodo.17141347>

Introduction:

In the current era of big data and artificial intelligence fitness trackers are getting more accurate as well as accepted and integrated into the daily lives of people. As such they offer new chances for personalized lifestyle guidance towards healthier habits. The Chatbot-based App VitaLo (DOCYET GmbH, Germany) integrates data from fitness trackers to help people improve lifestyle habits (sleep, physical activity etc.) through motivational interviewing and personalized lifestyle recommendations. In the conducted study the acceptance and usability of VitaLo as well as the effects on the participants lifestyle, especially in terms of physical activity, is analyzed.

Methods:

153 healthy adults were recruited for an observational, single-arm, prospective study with a duration of 30 days per participant, in which they were using both the Chatbot VitaLo and the fitness tracker Garmin Vivomove Sport. Changes in physical activity, sleep quality and well-being are evaluated with data tracked by the wearable and validated questionnaires. Furthermore, the general acceptance and usability are assessed.

Results:

The study is currently ongoing and expected to terminate in the beginning of May. The analysis will be conducted after study termination.

Discussion:

This study helps to better understand the potential and acceptance of chatbots as lifestyle advisors in combination with wearable data.

Kombinierte Versorgung als Konzept der digitalen Reha-Nachsorge: Ergebnisse der Patientenzufriedenheit und Benutzerfreundlichkeit

S. Knapp², F. Merten^{1,2*}, K. Kutz-Cerry², J. Huss²

1 f.merten@caspar-health.com, Caspar Clinic. Digitales Centrum für Gesundheit., Deutschland

2 Caspar Clinic. Digitales Centrum für Gesundheit., Deutschland

DOI: <https://doi.org/10.5281/zenodo.17141391>

Einleitung:

Die kombinierte Versorgung, als Umsetzungskonzept der digitalen Reha-Nachsorge, kombiniert die mit Software digital unterstützte Therapie mit persönlicher, therapeutischer Betreuung. Analysiert wurden daher Aspekte der Patientenzufriedenheit mit der therapeutischen Betreuung sowie der Benutzerfreundlichkeit einer Nachsorge-Applikation.

Methoden:

Grundlage der explorativen Datenanalyse (12.2024 - 02.2025) bildete eine Pilotierung eines PROM-Sets (u.a. ZUF-8, SUS). Dazu wurde u.a. zum Ende der Nachsorge ein digitaler Fragebogen erhoben. Acht Wochen nach Nachsorgestart wurde die Betreuungszufriedenheit mit dem CSAT erfragt. Für die Auswertungen wurden deskriptive Methoden sowie Korrelationen nach Pearson ($\alpha=0,05$) genutzt.

Ergebnisse:

Die Patienten waren im Durchschnitt 53,2 Jahre alt, zu 62,7% weiblich, 60,1% ganztags berufstätig und zu 56,1% aufgrund einer muskuloskelettalen Diagnose in Reha. Die Qualität der digitalen Reha-Nachsorge wurde zu 93,9% mit "gut" bis "ausgezeichnet" bewertet. Zu 91,6% hat die Maßnahme den "meisten" bis "allen" Bedürfnissen entsprochen. Mit der erhaltenen Hilfe waren 96,6% "weitgehend" bis "sehr zufrieden". Im Großen und Ganzen waren 97,6% der Patienten mit der digitalen Reha-Nachsorge "weitgehend" bis "sehr zufrieden". Es zeigten sich konsistent moderate Korrelationen von 0,23 - 0,33 zwischen dem CSAT und den Items der Patientenzufriedenheit ($p<0,05$). Die Benutzerfreundlichkeit zeigte sich hoch. Für die Einfachheit der Benutzung ergab sich im Mittel einen Wert von 4,6 (0,8), für die Funktionsintegration einen Wert 4,6 (0,8) sowie für die Sicherheit im Umgang einen Wert von 4,7 (0,7).

Diskussion:

Die Aspekte der therapeutischen Betreuung sowie die Umsetzung der digitalen Anwendung wurden durch die Befragung fokussiert. Die hohe Patientenzufriedenheit und Benutzerfreundlichkeit stärken somit den Ansatz der kombinierten Versorgung.

Digitale Transformation im Versorgungskontext: Praxisnahe Unterstützung hausärztlicher Teams durch Digitalisierungslots:innen

Aileen Hommel^{1*}, Alessia Gatti¹, Anika Zimmer¹, Jessica Skalmowski¹, Benjamin Senst¹, Tim Holetzke¹, Thomas Schrader², Christine Holmberg¹

*präsentierende Autorin

1 aileen.hommel@mhb-fontane.de, Institut für Sozialmedizin und Epidemiologie, Medizinische Hochschule Brandenburg Theodor Fontane, Deutschland

2 Fachbereich Informatik und Medien, Technische Hochschule Brandenburg, Deutschland

DOI: <https://doi.org/10.5281/zenodo.17141423>

Einleitung: Die digitale Transformation der Primärversorgung ist zentraler Bestandteil einer zukunftsorientierten Gesundheitsversorgung. In Brandenburg stellt die Umsetzung digitaler Anwendungen aufgrund der heterogenen digitalen Reife und einer überwiegend älteren Hausärzteschaft eine Herausforderung dar [2, 3]. Digitale Gesundheitsanwendungen (DiGA), elektronische Patientenakten (ePA) sowie sichere Kommunikationswege (KIM) bieten Potenzial zur Verbesserung der Versorgungsqualität. Studien weisen jedoch auf Implementierungshürden und Akzeptanzprobleme hin, insbesondere sobald etablierte Praxisabläufe gestört oder administrative Aufgaben erhöht werden [1, 4, 5].

Methoden: Das Projekt untersucht die Wirksamkeit eines Unterstützungsansatzes, bei dem Digitalisierungslots:innen Praxen gezielt und nutzerzentriert dabei begleiten, digitale Technologien und Prozesse nachhaltig einzuführen. Die begleitende Evaluation dieser Proof-of-Concept-Studie erfolgt im Mixed-Methods-Design. Veränderungen der digitalen Reife und Akzeptanz sind zentrale Indikatoren.

Ergebnisse: Bis März 2025 nutzte rund ein Drittel der eingeschlossenen Praxen das Unterstützungsangebot. Zentrale Anliegen umfassten Einführung und Nutzung von ePA und KIM, IT-Sicherheit und digitale Abrechnungsprozesse. Erste Rückmeldungen bestätigen bekannte Herausforderungen, insbesondere im Hinblick auf technische Störungen und administrative Belastungen. Eine weiterführende Analyse dieser Befunde erfolgt im Rahmen eines Prä-Post-Vergleichs im zweiten Quartal 2025.

Diskussion: Digitalisierungslots:innen leisten einen praxisnahen Beitrag zur Umsetzung digitaler Anwendungen in der Primärversorgung. Zukünftige Maßnahmen sollten über technische Lösungen hinaus auch soziale und organisationale Bedingungen berücksichtigen. Digitale Versorgungsformen müssen in bestehende Teamstrukturen, Kommunikationsprozesse und Arbeitsabläufe integrierbar sein, um Akzeptanz und Wirkung zu entfalten. Dies ist essenziell, um

unter den Bedingungen begrenzter Ressourcen die Funktionalität und Zukunftsfähigkeit ambulanter Versorgung zu sichern.

Referenzen: [1] Holetzek et al. (2025), [2] Kassenärztliche Bundesvereinigung (2023), [3] KBV Praxisbarometer Digitalisierung (2024), [4] Neunaber und Meister (2023), [5] Ziebland et al. (2021).

Digital Health Spark - Igniting need-driven innovation in healthcare: A study on motivation, self-confidence and understanding in technology transfer and business creation

Constanze Cavalier^{1*}, Linea Schmidt¹, Philipp Stoffers^{1,2}, Ariel Dora Stern¹

*presenter

¹ constanze.cavalier@student.hpi.de, Hasso Plattner Institute Potsdam, Germany

² Charité - Universitätsmedizin Berlin, Department of Hepatology and Gastroenterology, Campus Virchow-Klinikum (CVK) and Campus Charité Mitte (CCM), Berlin, Germany

DOI: <https://doi.org/10.5281/zenodo.17141488>

Introduction: Translating novel digital health solutions into clinical practice is challenging due to a complex, multi-stakeholder healthcare environment and regulatory constraints. The two-week program *Digital Health Spark—Igniting Need-Driven Innovation in Healthcare* at the Hasso Plattner Institute was designed to bridge the gap between theoretical knowledge and practical application in healthcare innovation. Based on the Stanford Biodesign approach, the course emphasised need-driven innovation and Design Thinking, including immersions in clinical settings.

This study evaluated whether this interdisciplinary program enhanced self-assessed innovation competencies in digital medicine. In collaboration with the Vivantes hospitals in Berlin, the course engaged master's, PhD students, and postdocs from medicine, computer science, and business disciplines in clinical shadowing, problem identification, and solution development based on real-world healthcare needs.

Methods: Using an adapted version of the HEInnovate EPIC questionnaire, including items specific to healthcare delivery and medical devices, the study measures changes in participants' motivation, confidence, and understanding of entrepreneurship and digital health innovation. Surveys were conducted pre- and post-course.

Results: The course engaged 36 participants in 14 interdisciplinary teams across two cohorts. Preliminary results from the 2024 cohort indicate that hands-on, need-driven learning fostered deeper understanding of healthcare challenges and stimulated creative ideation. The 2025 course iteration introduces enhanced validation phases to deepen innovation outcomes, survey evaluation is ongoing.

Discussion: This study provides early evidence that interdisciplinary, clinically-immersive education can strengthen entrepreneurial potential among participants and support early-stage innovation in academic environments.

Mehr Likes als Wirkung? Nutzung digitaler psychologischer Unterstützungsangebote bei Jugendlichen mit psychischen Störungen

S. H. Kohl^{1,2}, R. Krott¹, I.-K. Petras¹, A. T. Henn¹, Y. Terhorst³, K. Konrad^{1,2}, M. Siniatchkin¹

¹ simon.h.kohl@gmail.com Klinik für Psychiatrie, Psychosomatik und Psychotherapie des Kindes- und Jugendalters, Universitätsklinik RWTH Aachen, 52074 Aachen, Germany

² JARA-Institute Molecular Neuroscience and Neuroimaging (INM-11), Forschungszentrum Jülich, 52425 Jülich, Germany

³ Psychology Department, Ludwig-Maximilians-Universität München, 80802 München, Germany*

DOI: <https://doi.org/10.5281/zenodo.17141516>

Abstract

Einleitung:

Digitale psychologische Unterstützungsangebote (DPUs) umfassen digitale Formate – von sozialen Medien und Informationsplattformen bis hin zu strukturierten Interventionen und regulierten Angeboten wie Digitalen Gesundheitsanwendungen –, die im Zusammenhang mit psychischer Gesundheit genutzt werden können. Insbesondere unregulierte Angebote sind für Jugendliche mit psychischen Problemen leicht zugänglich und weit verbreitet. Dennoch ist bislang wenig darüber bekannt, wie DPUs tatsächlich genutzt und bewertet werden. Ziel der Studie ist die Entwicklung und erste Erprobung eines Fragebogens, der Nutzung, Einstellungen und Barrieren im Umgang mit DPUs systematisch erfasst.

Methoden:

Befragt werden Jugendliche (12–18 Jahre) aus der (teil-)stationären Kinder- und Jugendpsychiatrie der Uniklinik RWTH Aachen über eine anonyme Online-Plattform. Neben Angaben zur DPU-Nutzung werden demografische Merkmale, psychische Beschwerden, digitale Gesundheitskompetenz und Medienverhalten erhoben. Die Datenerhebung läuft und wird im Juli 2025 abgeschlossen.

Ergebnisse:

Vorläufige Auswertungen zeigen eine intensive Nutzung sozialer Medien (TikTok, Instagram, YouTube) und ChatGPT zur Bewältigung psychischer Belastungen, während evidenzbasierte Angebote kaum genutzt werden. Viele Jugendliche berichten von negativen Erfahrungen, etwa durch belastende Inhalte oder problematische Influencer:innen.

Diskussion:

Die Ergebnisse verdeutlichen eine Lücke zwischen hohem Bedarf und geringer Nutzung qualitätsgesicherter DPUs. Die Studie liefert empirische Grundlagen zur Untersuchung von Nutzungsmustern, Barrieren und Dissemination evidenzbasierter Angebote und schafft die Basis für weiterführende repräsentative Studien zur Erfassung der digitalen Gesundheitsversorgung Jugendlicher in Deutschland.

Effects of an app-based physical training for long and post COVID treatment (A PILOT study) on physical performance and psychological health

Sarah I. Krieg^{*1}, Lynn Matits^{1,2}, Lennart Mentz¹, Jana Schellenberg¹, Andreas Nieß³, Birgit Friedmann-Bette⁴, Peter Deibert⁵, Philipp Maier⁵, Michèle Natale⁴, Daniel A. Bizjak¹, Johannes Kirsten¹

*presenter (also underline presenter please)

1 Sports and Rehabilitation Medicine, University Hospital Ulm, Ulm, Germany,

2 Clinical & Biological Psychology, Institute of Psychology and Education, Ulm University, Ulm, Germany

3 Department of Sports Medicine, University Hospital Tübingen, Tübingen, Germany

4 Department of Sports Medicine, University Hospital Heidelberg, Heidelberg, Germany

5 Institute for Exercise and Occupational Medicine, Medical Centre and Faculty of Medicine, Albert-Ludwig-University, Freiburg, Germany

DOI: <https://doi.org/10.5281/zenodo.17141600>

Introduction: After SARS-CoV-2 infection, health-related quality of life (QoL) and physical performance are often reduced due to prolonged symptoms. Telerehabilitation can improve QoL in patients with chronic mental and somatic disorders as previous intervention studies show, suggesting potential benefits for the treatment of post COVID patients. We developed an app-based, multimodal physical training intervention tailored to long and post COVID symptoms. The objective is to evaluate its effects on physical performance, neurocognitive function, and QoL. We aim to reach patients who lack access to alternative treatment options, particularly due to infrastructural limitations.

Methods: A PILOT study is a prospective, randomized, two-armed multicenter study, aiming to include 100 patients with persistent or newly occurring symptoms related to a SARS-CoV-2 infection and suspected post COVID syndrome. Patients are randomized into an intervention group (n=50), receiving app-based endurance and strength training as well as progressive muscle relaxation and breathing exercises over eight weeks, and a control group receiving usual care (n=50). Physical performance, psychological variables and cognitive performance are evaluated.

Results: The effectiveness of an app-based training of the intervention group in comparison to the control group is analyzed. We constitute the adherence of app usage during the intervention as well as its effect on a change in QoL and physical performance.

Discussion: An app-based training can add value for patients by complementing available traditional forms of therapy. The findings of the PILOT study will contribute to understand the feasibility and effectiveness of app-based training for post-COVID patients, potentially improving future digital rehabilitation approaches.

Status Quo der Digitalisierung in der Dermatologie – Ergebnisse einer nationalen Befragung

L. Henk^{*1,2}, L. Schirren², V. Olsavszky³, F. Mühlensiepen⁴, C. Traidl-Hoffmann^{5,6}, M. Augustin⁷, A. Zink⁸, M. Hertl², S. Kuhn¹, J. Knitza¹

*präsentierender Autor (bitte auch unterstreichen)

1 lukas.henk@uni-marburg.de, Institut für digitale Medizin, Universitätsklinikum Gießen und Marburg, Philipps-Universität Marburg, Deutschland

2 Klinik für Dermatologie und Allergologie, Universitätsklinikum Gießen und Marburg, Marburg, Deutschland

3 Klinik für Dermatologie, Venerologie und Allergologie, Universitätsklinikum Mannheim, Mannheim, Deutschland

4 Zentrum für Versorgungsforschung Brandenburg, Medizinische Hochschule Brandenburg Theodor Fontane, Neuruppin, Deutschland.

5 Institut für Umweltmedizin und integrative Gesundheit, Medizinische Fakultät, Universitätsklinikum Augsburg, Augsburg, Deutschland

6 CK-Care, Christine Kühne Center for Allergy Research and Education, Davos, Schweiz

7 Institut für Versorgungsforschung in der Dermatologie und bei Pflegeberufen, Universitätsklinikum Hamburg Eppendorf, Hamburg, Deutschland

8 Klinik für Dermatologie und Allergologie, Technische Universität München, München, Deutschland

DOI: <https://doi.org/10.5281/zenodo.17142166>

Einleitung:

Digitale Medizin gewinnt zunehmend an Bedeutung in der Patientenversorgung – insbesondere in der visuell geprägten Dermatologie. Diese Studie untersuchte die Nutzung, Kenntnis und Wahrnehmung von digitalen Gesundheitsanwendungen (DiGAs) sowie die zunehmende Digitalisierung unter Dermatolog*innen in Deutschland.

Methoden:

Die Datenerhebung erfolgte anhand eines standardisierten Fragebogens unter dermatologisch tätigen Ärzt*innen während des Derma-Update-Kongress 2024 sowie Fortbildungsveranstaltungen der Uniklinik Marburg.

Ergebnisse:

Insgesamt nahmen 100 Ärzt*innen an der Umfrage teil (Durchschnittsalter 40,6 Jahre; 64% Frauen). Die selbst eingeschätzte Kenntnisse im Bereich der digitalen Medizin lag im mittleren Bereich (durchschnittlich 3,2/5). 42% der Ärzt*innen gaben an Patient*innen auch telemedizinisch zu versorgen. Das eigene Wissen zu DiGAs wurde als gering eingestuft (durchschnittlich 3,9/10), 21% verordneten bereits DiGAs, wobei 98% der Teilnehmenden angaben, dass DiGAs die medizinische Versorgung verbessern können. Fehlende Kenntnisse zur Implementierung (76%) und Zeitmangel (74%) wurden als Hauptbarrieren gesehen. Der allgemeine Einfluss der Digitalisierung auf die Patient*innenversorgung wurde positiv gesehen (durchschnittlich 4,1/5), der Wunsch nach Schulungen und Fortbildungen auf dem Gebiet war groß (80%).

Diskussion:

Obwohl digitale Medizin, Telemedizin und DiGAs überwiegend positiv bewertet werden, ist ihre Anwendung im dermatologischen Alltag noch gering. Spezifische Schulungsangebote stellen einen wesentlichen Hebel zur Förderung der Implementierung dar.

Digitalisierung im Sozial- und Gesundheitswesen: Zwischen Offenheit und strukturellen Hürden

A. Wiebe¹, M. Kappe^{1*}, M. Jurisch², W. Swoboda¹, J. Schobel¹

*präsentierender Autor

1 miriam.kappe@hnu.de, DigiHealth Institut, Hochschule für angewandte Wissenschaften Neu-Ulm, Deutschland

2 AristaFlow GmbH, Deutschland

DOI: <https://doi.org/10.5281/zenodo.17142291>

Einleitung: Digitalisierung wird als zentraler Lösungsansatz zur Bewältigung aktueller Herausforderungen im Sozial- und Gesundheitswesen (z.B. Fachkräftemangel, demographischer Wandel) diskutiert. Ihrem Fortschritt stehen jedoch strukturelle und organisatorische Hürden entgegen. Diese Interviewstudie ist Teil des Forschungsprojektes KompIGA und untersucht den aktuellen Stand der Digitalisierung im Sozial- und Gesundheitswesen. Sie stellt einen ersten Schritt zur systematischen Erhebung digitaler Transformationsprozesse innerhalb der Projektregion dar.

Methoden: Im Rahmen der Studie wurden 6 leitfadengestützte Interviews mit Personen durchgeführt, die aktiv an Digitalisierungsprojekten in der Domäne arbeiten. Die Auswertung erfolgte anhand der qualitativen Inhaltsanalyse nach Kuckartz.

Ergebnisse: Die Befragten beschreiben sowohl ihre Organisationen als auch das Sozial- und Gesundheitswesen insgesamt als schwach digitalisiert. Sie betonen eine zunehmende Offenheit gegenüber digitalen Lösungen, sehen jedoch die Fortschritte als gering. Hauptverantwortlich sind fehlende finanzielle Mittel. Zudem wird häufig versäumt, Ressourcen für die grundlegende IT-Infrastruktur einzuplanen. Verbesserungsbedarf wird besonders im administrativen Bereich und der Planung gesehen – weniger in der direkten Pflege. Konzepte zum Umgang mit größeren IT-Ausfällen oder Cyberangriffen befinden sich in der Entwicklung. Digitale Kompetenzen der Mitarbeitenden werden als sehr heterogen wahrgenommen, bislang fehlen systematische Weiterbildungsangebote.

Limitationen: Die geringe Anzahl an Teilnehmenden schränkt die Aussagekraft der Studie deutlich ein. Eine Erweiterung der Stichprobe und weitere Studien zur Erfassung des Stands der Digitalisierung können belastbarere Ergebnisse liefern.

Diskussion: Die Ergebnisse zeigen, dass trotz erkennbarer Fortschritte weiterhin strukturelle Hürden die Digitalisierung im Sozial- und Gesundheitswesen bremsen. Zwar ist die Offenheit gegenüber digitalen Lösungen gestiegen, das wahrgenommene Digitalisierungspotenzial ist aber weiterhin hoch. Insgesamt besteht umfassender Handlungsbedarf, um Digitalisierung nachhaltig und praxisnah zu verankern.

Treating Metabolic Syndrome and Steatotic Liver Disease: A Systematic Literature Review of Conventional, Digital, and Personalized Just-in-Time Treatment Interventions

E. Habib^{1*}, D. Kaya^{1,6 *}, P. Jeya Subha Jeyaraj Samraj^{1,5}, E. Rosner¹, C. Scheuringer¹, L. Spies¹, A. Stern¹, K. Mai^{2,3,5}, R. Paprott¹, S. Konigorski¹, P. Stoffers^{1,4}

*(eva.habib@student.hpi.uni-potsdam.de, dilem.kaya2@student.hpi.uni-potsdam.de) presenters

1 Hasso-Plattner Institute, University of Potsdam, Germany1

2 Department of Endocrinology and Metabolism, Charité - Universitätsmedizin Berlin, Germany2

3 Center for Cardiovascular Research (CCR), Charité - Universitätsmedizin Berlin, Germany3

4 Charité - Universitätsmedizin Berlin, Department of Hepatology and Gastroenterology, Campus Virchow-Klinikum (CVK) and Campus Charité Mitte (CCM), Berlin, Germany

5 Berlin Institute of Health (BIH), Charité - Universitätsmedizin Berlin, Germany7

6 SAP SE, Potsdam, Germany1

DOI: <https://doi.org/10.5281/zenodo.17142331>

Introduction:

Metabolic dysfunction-associated steatotic liver disease (MASLD) is the hepatic manifestation of metabolic syndrome (MetS), a condition characterized by at least three of five NCEP ATP III criteria: elevated waist circumference, hypertension, reduced high-density lipoprotein, elevated fasting glucose, and elevated triglycerides. MASLD has become the most prevalent chronic liver disease globally, affecting 30% of adults with a 1% annual increase². Both conditions are strongly influenced by modifiable lifestyle factors, highlighting the opportunity for personalized digital therapeutic approaches.

Methods:

To better understand conventional treatment approaches and how they can be (and already are) translated into digitally supported treatments a systematic search following the PRISMA guidelines will be conducted in PubMed, IEEE Xplore, ACM Digital Library and DIMDI. Eligible studies will include clinical trials, pilot studies, randomized controlled trials and meta-analyses published in English or German from 2015 to April 2025 that evaluate interventions targeting any component of the MetS and/or MASLD using any intervention modality in adult populations.

Results:

Results will be organized according to NCEP ATP III criteria. Two independent reviewers will extract data using a standardized template, with discrepancies resolved through consultation with a third reviewer. Our preliminary search using specified terms (see appendix) indicates extensive

knowledge of interventions for MetS and/or MASLD, yet reveals that only a small fraction has been implemented in mobile health applications.

Discussion:

By systematically analyzing the gap between evidence-based interventions and currently available mobile health offerings, this study aims to inform the development of innovative digital solutions that integrate personalized approaches to metabolic health management.

Literature

1. Henry, L., Paik, J. & Younossi, Z. M. (2022). Review article: the epidemiologic burden of non-alcoholic fatty liver disease across the world. *Alimentary Pharmacology & Therapeutics*, 56(6), 942–956. <https://doi.org/10.1111/apt.17158>
2. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). (2001). *ATP III guidelines at-a-glance quick desk reference*. National Heart, Lung, and Blood Institute, National Institutes of Health.

Kombinierte Versorgung als Konzept der digitalen Reha-Nachsorge:

Ergebnisse der Patientenzufriedenheit und Benutzerfreundlichkeit

S. Knapp², F. Merten^{1,2*}, K. Kutz-Cerry², J. Huss²

1 f.merten@caspar-health.com, Caspar Clinic. Digitales Centrum für Gesundheit., Deutschland

2 Caspar Clinic. Digitales Centrum für Gesundheit., Deutschland

Einleitung:

Die kombinierte Versorgung, als Umsetzungskonzept der digitalen Reha-Nachsorge, kombiniert die mit Software digital unterstützte Therapie mit persönlicher, therapeutischer Betreuung. Analysiert wurden daher Aspekte der Patientenzufriedenheit mit der therapeutischen Betreuung sowie der Benutzerfreundlichkeit einer Nachsorge-Applikation.

Methoden:

Grundlage der explorativen Datenanalyse (12.2024 - 02.2025) bildete eine Pilotierung eines PROM-Sets (u.a. ZUF-8, SUS). Dazu wurde u.a. zum Ende der Nachsorge ein digitaler Fragebogen erhoben. Acht Wochen nach Nachsorgestart wurde die Betreuungszufriedenheit mit dem CSAT erfragt. Für die Auswertungen wurden deskriptive Methoden sowie Korrelationen nach Pearson ($\alpha=0,05$) genutzt.

Ergebnisse:

Die Patienten waren im Durchschnitt 53,2 Jahre alt, zu 62,7% weiblich, 60,1% ganztags berufstätig und zu 56,1% aufgrund einer muskuloskelettalen Diagnose in Reha. Die Qualität der digitalen Reha-Nachsorge wurde zu 93,9% mit “gut” bis “ausgezeichnet”

bewertet. Zu 91,6% hat die Maßnahme den “meisten” bis “allen” Bedürfnissen entsprochen. Mit der erhaltenen Hilfe waren 96,6% “weitgehend” bis “sehr zufrieden”. Im Großen und Ganzen waren 97,6% der Patienten mit der digitalen Reha-Nachsorge “weitgehend” bis “sehr zufrieden”. Es zeigten sich konsistent moderate Korrelationen von 0,23 - 0,33 zwischen dem CSAT und den Items der Patientenzufriedenheit ($p<0,05$). Die Benutzerfreundlichkeit zeigte sich hoch. Für die Einfachheit der Benutzung ergab sich im Mittel einen Wert von 4,6 (0,8), für die Funktionsintegration einen Wert 4,6 (0,8) sowie für die Sicherheit im Umgang einen Wert von 4,7 (0,7).

Diskussion:

Die Aspekte der therapeutischen Betreuung sowie die Umsetzung der digitalen Anwendung wurden durch die Befragung fokussiert. Die hohe Patientenzufriedenheit und Benutzerfreundlichkeit stärken somit den Ansatz der kombinierten Versorgung.

Strukturierte Lehre der Grundlagen von Digitaler Medizin im Medizinstudium in Form von digitalen Lerneinheiten in Articulate Rise

M. Brand¹, P. Lantwin¹*, N. Kunz¹, L. Sprehe¹, P. Knaup¹

1 petra.knaup@med.uni-heidelberg.de, Institut für Medizinische Informatik, Heidelberg, Deutschland

*präsentierende Autorin

DOI: <https://doi.org/10.5281/zenodo.17142437>

Einleitung:

Damit in der heutigen Zeit der zunehmenden Digitalisierung die Patientenversorgung sicherer und effektiver gestaltet werden kann, müssen Ärzt*innen digitale Kompetenzen erwerben¹⁻³. Im Medizinstudium der Medizinischen Fakultät Heidelberg (MFHD) ist Digitale Medizin an verschiedenen Stellen präsent, eine strukturierte und einheitliche Vermittlung der Grundlagen findet zurzeit erst gegen Ende des klinischen Abschnitts statt. In Vorbereitung auf die neue Approbationsordnung fördert die MFHD die Entwicklung eines neuen Lehrkonzepts für digitale Kompetenzen (Projekt HeiCuDigiMed).

Methoden:

Es wurden Veranstaltungen zu Themen der Digitalen Medizin im Heidelberger Medizinstudium identifiziert. Auf Basis dieser Informationen, sowie den Lernzielen des NKLM-Projektgruppenkatalogs „Digitale Kompetenzen“, wurden Themen für Grundlagen-orientierte digitale Lerneinheiten systematisch ausgewählt⁴. Die Lerneinheiten werden von einem Team aus Ärzt*innen und Medizininformatiker*innen als Microlearnings mit Articulate Rise konzipiert und in einen Moodle-Kurs integriert.

Ergebnisse:

Die Grundlagen der folgenden Themen werden in Form jeweils einer digitalen Lerneinheit aufgearbeitet: Digitalisierung, Big Data, Apps, (elektronische) Patientenakte, E-Health, Telemedizin, Krankenhausinformationssysteme, Künstliche Intelligenz und Bildverarbeitung. Die Lerneinheiten sind kompakt und sollen bereits in der Vorklinik bereitgestellt werden. Im Moodle-Kurs sind auch weiterführende Materialien, beispielsweise die Lernplattform für Künstliche Intelligenz „KI-Campus“, verlinkt⁵.

Diskussion:

Mit digitalen Lerneinheiten zu Grundlagen der Digitalen Medizin legen wir eine terminologische und inhaltliche Basis für Studierende und Lehrende der MFHD. In allen Fächern kann auf diese zentralen Konzepte zurückgegriffen werden und neue Lehrveranstaltungen können darauf aufbauen. Die Inhalte sollen erstmals im Wintersemester 2025/26 zur Verfügung stehen und regelmäßig durch die Teilnehmenden evaluiert werden. Schlüsselinhalt werden formativ online in den Lerneinheiten überprüft und sind Teil des summativen Assessments in einer Klausur.

Literatur:

1. Sahan F, Guthardt L, Panitz K, et al. Enhancing Digital Health Awareness and mHealth Competencies in Medical Education: Proof-of-Concept Study and Summative Process Evaluation of a Quality Improvement Project. *JMIR Med Educ.* 2024;10:e59454. doi:10.2196/59454
2. Digitalisierung H. The Digital Turn. *Hochschulbildung Im Digit Zeitalt Arbeitspapier Nr 27 Berl Hochschulforum Digit Verfügbar Unter Httpshochschulforumdigitalisierung DesitesdefaultfilesdateienAbschlussbericht Pdf.* Published online 2016.
3. Pisarenco C, Pisarenco S. Digitalisation of Health Data: Impact on Patient-Centredness and Safety. *Bul Acad Ştiinţe Mold Ştiinţe Medicale.* 2024;80(3):139-143.
4. Medizinischer Fakultätentag. Nationaler Kompetenzbasierter Lernzielkatalog Medizin (NKLM), Projektgruppenkatalog Digitale Kompetenzen. April 27, 2021. Accessed April 7, 2025. <https://nkml.de/zend/objective/list/orderBy/@objectivePosition/studiengang/Themen/zeitsemester/Projektgruppenkataloge/fachsemester/Digitale%20Kompetenzen>
5. *Digitalisierung in Studium und Lehre gemeinsam gestalten: Innovative Formate, Strategien und Netzwerke.* 1st ed. 2021. Springer Fachmedien Wiesbaden; 2021.

Internationale Strategien zur Verbesserung der Interoperabilität im Austausch von Gesundheitsinformationen: Ein Scoping Review

A-L. Brecher^{1*}, L. Kraft¹, A. E. Amelung¹

*presenter

1 brecher.anna-lena@mh-hannover.de, Medizinische Hochschule Hannover, Institut für Epidemiologie, Sozialmedizin und Gesundheitssystemforschung, Deutschland

DOI: <https://doi.org/10.5281/zenodo.17142459>

Einleitung:

Der reibungslose Austausch digitaler Gesundheitsdaten zwischen verschiedenen Akteuren, Sektoren und Anwendungen ist eine zentrale Voraussetzung für die erfolgreiche digitale Transformation des Gesundheitswesens. Im deutschen Gesundheitswesen bestehen derzeit jedoch noch erhebliche Defizite in Bezug auf offene Schnittstellen und einheitliche Standards. Ziel der Studie ist es daher, einen umfassenden Überblick über internationale Strategien und Konzepte zur Verbesserung der Interoperabilität im Gesundheitswesen zu gewinnen und übertragbare Best Practices für das deutsche Gesundheitssystem zu identifizieren.

Methoden:

Das Scoping Review wird gemäß der JBI-Methodik und unter Berücksichtigung der PRISMA-Checkliste durchgeführt. Die systematische Literatursuche erfolgt in den Datenbanken PubMed, Web of Science und CINAHL. Ergänzend wird eine manuelle Suche durchgeführt, um weitere relevante Publikationen sowie graue Literatur einzubeziehen. Die Auswahl der Studien findet in einem zwei-stufigen Screening-Prozess durch zwei unabhängige Reviewer statt. Anschließend werden die relevanten Daten extrahiert, systematisch analysiert und zentrale Erfolgsfaktoren auf ihre Übertragbarkeit auf das deutsche Gesundheitssystem diskutiert.

Ergebnisse:

Die Durchführung des Scoping Reviews wird im Juni 2025 abgeschlossen, so dass die Ergebnisse im September 2025 vorgestellt werden können. Der Fokus liegt dabei auf internationalen Best Practices - z.B. in Bezug auf Governance-Strukturen, politische Strategien und Standardisierungsprozesse - sowie auf der Diskussion von Handlungsoptionen für das deutsche Gesundheitssystem.

Diskussion:

Die systematische Analyse der internationalen Literatur zu interoperabilitätsfördernden Strategien und Konzepten im Kontext des Datenaustauschs im Gesundheitswesen liefert einen systematischen Überblick über den aktuellen Stand. Darauf aufbauend können Handlungsempfehlungen für das deutsche Gesundheitssystem abgeleitet werden, um die

Interoperabilität gezielt zu verbessern und die Implementierung innovativer, sektorenübergreifender Gesundheitslösungen nachhaltig zu stärken.

A Compliant Mechanism Based Prosthetic Foot for LMICs

Kanta Sarma^{1*}, Darsh Nilesh Shah²,

¹ kanta.sarma@st.ovgu.de, Otto-von-Guericke-University Magdeburg, Germany

² darsh.shah@st.ovgu.de, Otto-von-Guericke-University Magdeburg, Germany

DOI: <https://doi.org/10.5281/zenodo.17142579>

Introduction: Over 80% of individuals with lower-limb amputations, amounting to more than 30 million people, reside in low- and middle-income countries (LMICs)[3]. Only 10% of the amputee who have access that too with advanced healthcare system can only afford the existing biomechanical prostheses of the market. Apart from the cost, these devices come with limitations too. Such as implant loosening, wear induced osteolysis, loss of natural kinematics. Consequently, that leads rest of the 90% opting for other inconvenient devices like crutch, limbs, dysfunctional prosthesis etcetera.

Methods: Functional and affordable access to prosthetic feet remains a significant challenge in most LMICs due to the high cost of traditionally designed biomechanical prostheses. The following study will outline the design process of a compliant mechanism-based prosthetic foot, which was designed to be an affordable, durable, and functional alternative that also conforms to international standards of both ISO 10328.

Results: The design included an ESR structure, using a multi-keel structure, inspired by leaf springs, for biomechanical efficiency and natural gait cycle mimicry. Material selection focused on Nylon 6 for strength, flexibility, and reasonable cost. Finite element simulations were conducted to evaluate mechanical performance, and results indicated areas for further refinement in fatigue resistance.

Discussion: This research is part of the current efforts to produce low-cost, high-performance prosthetics with improved mobility and quality of life for amputees in resource-poor settings.

Digitale Kluft statt digitaler Teilhabe? Sozioökonomische Ungleichheiten in der digitalen Gesundheitskompetenz

Lara Reinke^{* 1}, Ilka Budde², Konrad Fuchs³, Alexander Hann⁴, Matthias Kiesel⁵, Andreas S. Kunz⁶, Matthias Nieberler⁷, Kristen Rak⁸, Florian P. Reiter⁹, Daniel Röder¹⁰, Rüdiger Pryss^{11, 12}, Marc Schmalzing¹³, Astrid Schmieder¹⁴, Johannes

Schrenker ¹⁵, August Stich ¹⁶, Stefan Störk^{17, 18}, Patrick-Pascal Strunz ¹³, Fabian Kerwagen ^{17, 18} für die iAGDM Forschungsgruppe

1 lara.reinke@stud-mail.uni-wuerzburg.de, Medizinische Fakultät der Universität Würzburg

2 Medizinische Klinik und Poliklinik I, Endokrinologie und Diabetologie, Universitätsklinikum Würzburg

3 Klinik und Poliklinik für Unfall-, Hand-, Plastische und Wiederherstellungschirurgie, Universitätsklinikum Würzburg

4 Medizinische Klinik und Poliklinik II, Gastroenterologie, Universitätsklinikum Würzburg

5 Frauenklinik und Poliklinik, Universitätsklinikum Würzburg

6 Institut für Diagnostische und Interventionelle Radiologie, Universitätsklinikum Würzburg

7 Klinik und Poliklinik für Psychiatrie, Psychosomatik und Psychotherapie, Universitätsklinikum Würzburg

8 Klinik und Poliklinik für Hals-, Nasen- und Ohrenheilkunde, Kopf- und Hals-Chirurgie, Universitätsklinikum Würzburg

9 Medizinische Klinik und Poliklinik II, Hepatologie, Universitätsklinikum Würzburg

10 Klinik und Poliklinik für Anästhesiologie, Intensivmedizin, Notfallmedizin und Schmerztherapie, Universitätsklinikum Würzburg

11 Institut für klinische Epidemiologie und Biometrie, Universität Würzburg, Würzburg

12 Institut für medizinische Datenwissenschaften, Universitätsklinikum Würzburg, Würzburg

13 Medizinische Klinik II, Rheumatologie/Klinische Immunologie, Universitätsklinikum Würzburg

14 Klinik für Dermatologie und Venerologie, Universitätsklinikum Würzburg

15 Abteilung für Zahnärztliche Prothetik, Universität Würzburg

16 Medizinische Klinik und Poliklinik II, Infektiologie/Tropenmedizin, Universitätsklinikum Würzburg

17 Medizinische Klinik und Poliklinik I, Kardiologie, Universitätsklinikum Würzburg

18 Deutsches Zentrum für Herzinsuffizienz, Department für klinische Forschung und Epidemiologie, Universitätsklinikum Würzburg

DOI: <https://doi.org/10.5281/zenodo.17142648>

Hintergrund: Die digitale Transformation der Medizin erfordert digitale Gesundheitskompetenz (dGK) der Patient:innen. Gesetzliche Krankenkassen (GKV) sind verpflichtet, entsprechende Förderangebote bereitzustellen. Ausprägung und soziale Determinanten der dGK sowie die Bekanntheit von GKV-Schulungsangeboten sind bislang unzureichend erforscht.

Methoden: Von Juli 2024 bis April 2025 wurde eine anonymisierte Querschnittsbefragung in 14 Ambulanzen des Universitätsklinikums Würzburg durchgeführt (Abbildung 1). Erhoben wurden

soziodemografische Merkmale, die dGK mittels GR-eHEALS-Fragebogen (min. 1; max. 5), Nutzung digitaler Gesundheitsanwendungen sowie Wissen über und Einstellungen zu dGK-Schulungen.

Ergebnisse: 1.643 Patient:innen nahmen teil (53 ± 18 Jahre; 46 % Frauen). 88 % nutzten bereits mindestens ein digitales Gesundheitsprodukt, meist das eRezept. Die mittlere dGK lag bei $3,6 \pm 0,9$, wobei Patient:innen mit Hauptschulabschluss ($3,2 \pm 1,1$) und niedrigem monatlichem Einkommen (<2.000 €, $3,3 \pm 1,1$) eine geringere dGK als Teilnehmende mit Abitur ($3,9 \pm 0,8$) oder hohem Einkommen (>3.500 €, $3,9 \pm 0,7$) aufwiesen (Abbildung 2A). Hauptschulabschluss (OR 3,6; 95% KI 2,6-4,9) und niedriges Einkommen (OR 2,9; 95% KI 2,1-4,0) waren alters- und geschlechtsunabhängige Prädiktoren für eine niedrige dGK (Abbildung 2B). Schulungsangebote zur dGK waren 31 % der Befragten bekannt und 16 % hatten bislang ein solches Angebot wahrgenommen, wobei die Inanspruchnahme bei bildungs- (Hauptschule: 13 % vs. Abitur: 19 %) und einkommensschwachen Gruppen (<2.000 €: 13 % vs. >3.500 €: 19 %) geringer war. Die grundsätzliche Teilnahmebereitschaft an dGK-Schulungen war mit 68 % hoch.

Schlussfolgerung: Die dGK ist ungleich verteilt. Gruppen mit geringem sozioökonomischem Status zeigten einen erhöhten Unterstützungsbedarf, waren jedoch besonders schlecht über bestehende Angebote informiert. Eine gezieltere Unterstützung und barrierearme Gestaltung von Schulungsmaßnahmen sind erforderlich, um digitale Gesundheitsteilhabe chancengleich zu ermöglichen.

Verwendung von Fitnesstrackern zur kontinuierlichen Vitalparameterüberwachung bei postoperativen Patienten

P. Helmer^{1*}, J. Brugger¹, R. Pryss², P. Meybohm¹, M. Sammeth¹

*präsentierender Autor

1 helmer_p@ukw.de, Klinik und Poliklinik für Anästhesiologie, Intensivmedizin, Notfallmedizin und Schmerztherapie, Universitätsklinikum Würzburg, DE

2 Lehrstuhl für Klinische Epidemiologie und Biometrie, Universitätsklinikum Würzburg, DE

DOI: <https://doi.org/10.5281/zenodo.17142705>

Einleitung:

Die kontinuierliche Überwachung von Vitalparametern bei postoperativen Patienten könnte aufgrund der frühzeitigen Diagnostik postoperativer Komplikationen zukünftig die Patientensicherheit erhöhen. Technische und personelle Einschränkungen limitieren derzeit jedoch eine flächendeckende Umsetzung solcher Systeme. Fitnesstracker könnten aufgrund ihrer breiten Verfügbarkeit und Akzeptanz eine praktikable Alternative darstellen. Daher stellt sich die Frage, ob diese Geräte geeignet sind, um zukünftig die postoperative Überwachung von Patienten zu ergänzen.

Methoden:

Ziel dieses Projekts war die Untersuchung der Messgenauigkeit von Vitalparametern durch Fitnesstracker im Vergleich zu etablierten medizinischen Standardgeräten. In zwei prospektiven klinischen Studien wurden handelsübliche Fitnesstracker (Apple Watch 7, Garmin Fenix 6 Pro, Withings Scanwatch) im Aufwachraum ($n=201$) sowie auf einer Überwachungsstation ($n=36$) mit etabliertem Klinikstandard (Philips MX750/IntelliVueX3 inklusive 3-Kanal-EKG, Pulsoxymetrie und Impedanzpneumographie) zur Überwachung der Herzfrequenz (HF), Sauerstoffsättigung (SpO₂) und Atemfrequenz (RR) verglichen. Zusätzlich wurde die Patientenzufriedenheit, Akzeptanz und Einstellung der Patienten gegenüber einer kontinuierlichen Vitalparameterüberwachung erfasst.

Ergebnisse:

Alle untersuchten Fitnesstracker zeigten bei der HF hohe Korrelation ($r=0,95-0,98$; $p<0,001$), geringen Bias (-1,2 bis 0,36 /min) und geringe Ausfallrate (0,02 %-2,7 %) gegenüber dem Standardmonitoring. SpO₂ Messungen wiesen hingegen nur schlechte bis moderate Korrelation ($r=0,21-0,64$; $p<0,001$) und hohe Dropout-Raten (13,8 %-48,6 %) auf, abhängig vom jeweiligen Hersteller. Die RR zeigte eine schlechte Korrelation ($r=0,28$; $p<0,001$). Patienten äußerten hohe

Zufriedenheit, empfanden das Tragen der Geräte als angenehm und waren mehrheitlich bereit, ihre Daten für Forschungszwecke bereitzustellen.

Diskussion:

Fitnesstracker könnten zukünftig eine ergänzende Rolle bei der kontinuierlichen Überwachung der HF bei postoperativen Patienten einnehmen, wohingegen die Messungen von SpO₂ und RR Raum für Optimierung bieten. Dennoch belegt das hohe Patienteninteresse das Potenzial solcher Technologien für zukünftige Forschungsansätze und patientenorientierte Überwachungsstrategien.

Addressing AI Bias Through Data Disclosure: Evidence from a randomized Experiment among Physicians

Louis Agha-Mir-Salim¹, Anna-Lena Fehlhaber², Nicolas Frey¹, Ivar Krumpal³, Anatol-Fiete Näher^{1,4}

1. Institute of Medical Informatics, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany
2. Institute of Practical Computer Science, Leibniz University Hannover, Hannover, Germany
3. Institute of Sociology, University of Leipzig, Leipzig, Germany
4. Digital Global Public Health, Hasso Plattner Institute for Digital Engineering, University of Potsdam, Potsdam, Germany

DOI: <https://doi.org/10.5281/zenodo.17142782>

Introduction: The integration of machine learning (ML) into healthcare offers far-reaching opportunities to improve health outcomes. However, concerns regarding algorithmic fairness persist, particularly when ML models are trained on datasets that underrepresent specific patient populations. Transparency regarding training data has been proposed as a strategy to address these challenges. This study examines how demographic representation in training data affects physician adoption of AI predictions.

Methods: A within-subjects factorial survey experiment was conducted online between March and May 2025 among physicians of all seniority levels. Participants evaluated eight vignettes describing hypothetical ML systems for predicting acute coronary syndrome. Each vignette systematically manipulated holdout data demographics (age and gender representation) while controlling for model type and dataset size, enabling causal inference on physician decision-making. Primary outcomes were perceived relevance (7-point Likert scale) and intention to use (yes/no).

Results: Inadequate representation of age and gender groups in holdout data reduced physicians' perceived relevance and intention to use ML systems. For every percentage point increase in age difference between holdout data and physician's patient population, relevance ratings decreased by 0.056 points ($p < 0.001$) and usage likelihood decreased by 1.5%, ($p < 0.001$). Similarly, inadequate representation of gender reduced both relevance ($\beta = -0.014$, $p < 0.001$) and use intention (0.41%, $p < 0.01$). The effects were consistent across physician characteristics, demonstrating that demographic representation in training data substantially affects clinical AI adoption decisions.

Discussion: This study contributes to evaluating how transparency about data composition influences physicians' perceptions and intended use of clinical ML systems. The findings help promoting appropriate use of ML in healthcare while meeting applicable ethical standards.