

Digitale Gesundheitsinterventionen entwickeln, evaluieren und implementieren (Teil I) – Diskussionspapier der Arbeitsgruppe Digital Health des Deutschen Netzwerk Versorgungsforschung (DNVF)

Development, Evaluation and Implementation of Digital Health Interventions (Part 1) – Discussion Paper of the Digital Health Working Group of the German Network for Health Services Research (DNVF)

Autorinnen/Autoren

Sven Kernebeck¹ , Madlen Scheibe², Monika Sinha^{3, 4}, Florian Fischer^{5, 6}, Andreas Knapp⁷ , Patrick Timpel^{2, 8}, Lorenz Harst² , Ulrich Reininghaus^{9, 10, 11}, Horst Christian Vollmar¹² 

Institute

- 1 Lehrstuhl für Didaktik und Bildungsforschung im Gesundheitswesen – Fakultät für Gesundheit, Universität Witten Herdecke, Witten, Germany
- 2 Zentrum für Evidenzbasierte Gesundheitsversorgung, Universitätsklinikum und Medizinische Fakultät Carl Gustav Carus an der TU Dresden, Dresden, Germany
- 3 Mitglied AG Bioinformatik, Charité Universitätsmedizin Berlin, Berlin, Germany
- 4 Beratung im Gesundheitswesen – angewandte Versorgungsforschung, SINHA, Berlin, Germany
- 5 Bayerisches Forschungszentrum Pflege Digital, Hochschule Kempten, Kempten, Germany
- 6 Institut für Public Health, Charité Universitätsmedizin Berlin, Berlin
- 7 Mint Medical GmbH, Heidelberg
- 8 Wissenschaftliches Institut für Gesundheitsökonomie und Gesundheitssystemforschung, WIG2 GmbH, Leipzig, Germany
- 9 Department of Public Mental Health, Central Institute of Mental Health, University of Heidelberg, Mannheim, Germany
- 10 ESRC Centre for Society and Mental Health, King's College London, London, United Kingdom of Great Britain and Northern Ireland
- 11 Centre for Epidemiology and Public Health, Health Service and Population Research Department, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, Germany
- 12 Abteilung für Allgemeinmedizin (AM RUB), Medizinische Fakultät, Ruhr-Universität Bochum, Bochum, Germany

Schlüsselwörter

Digital Health, digitale Interventionen, Evaluation, Implementierung, Versorgungsforschung, E-Health

Key words

digital health, digital interventions, E-Health, evaluation, implementation, health services research

online publiziert 29.11.2022

Bibliografie

Gesundheitswesen 2023; 85: 58–64

DOI 10.1055/a-1933-2779

ISSN 0941-3790

© 2022. Thieme. All rights reserved.

Georg Thieme Verlag, Rüdigerstraße 14, 70469 Stuttgart, Germany

Korrespondenzadresse

Dr. Sven Kernebeck, MPH
Universität Witten Herdecke
Lehrstuhl für Didaktik und Bildungsforschung im Gesundheitswesen – Fakultät für Gesundheit
Alfred-Herrhausen-Straße 50
58448 Witten
Germany
Sven.kernebeck@uni-wh.de

ZUSAMMENFASSUNG

Digitale Interventionen (DI) zu gesundheitsbezogenen Themen gewinnen zunehmend an Bedeutung. Oftmals sind sie komplexe Interventionen, die in einem komplexen Versorgungskontext evaluiert und implementiert werden. Aufgrund der Merkmale digitaler Interventionen ist die Versorgungsforschung vor methodische Herausforderungen gestellt. Daher hat die Arbeitsgruppe Digital Health des Deutschen Netzwerks Versor-

gungsforschung (DNVF) ein Diskussionspapier erstellt. Darin wird diskutiert, welche Herausforderungen mit der Entwicklung und Evaluation von DI einhergehen. Es werden Lösungsansätze und zukünftige Forschungsbedarfe aufgezeigt.

ABSTRACT

The development and application of digital interventions in health-related topics are gaining momentum in health service research. Digital interventions are often complex and need to be evaluated and implemented in complex settings. Due to

their characteristics, this poses methodological challenges for health services research that have to be identified and addressed. Hence, the Working Group on Digital Health of the German Network for Health Services Research (DNVF) has prepared a discussion paper. This paper discusses methodological, practical and theoretical challenges associated with the development and evaluation of digital interventions from the perspective of health services research. Possible solutions are suggested and future research needs to address these methodological challenges are identified.

Einleitung

Die Entwicklung und Anwendung digitaler Interventionen (DI) zu gesundheitsbezogenen Themen gewinnen sowohl in der Forschung als auch in Politik und Gesellschaft zunehmend an Bedeutung [1]. Die Weltgesundheitsorganisation (WHO) definiert Digital Health Interventionen als "different ways in which digital and mobile technologies are being used to support health system needs." [2]. Die Anwendung von DI mit gesundheitlichem Bezug bzw. der Digitalisierung im Gesundheitswesen grundsätzlich, wird nicht als Selbstzweck verstanden, sondern sie dient dem Wohl der Patient*innen und Bürger*innen [3].

Aus Perspektive der Versorgungsforschung ergeben sich verschiedene Herausforderungen im Kontext von DI: Erstens sind DI ein relevanter Gegenstand der Versorgungsforschung, da sie beispielsweise als komplexe Intervention für verschiedene Krankheitsbilder entwickelt, evaluiert und implementiert werden [4, 5]. Zweitens stellen DI eine methodische, epistemologische und wissenschafts-ethische Herausforderung für die Evaluation – und damit einem der wichtigsten Betätigungsfelder der Versorgungsforschung – dar. Sie erfordern neue Analysemethoden, Datentypen und -quellen [4, 5].

Diese und weitere Herausforderungen bei der Evaluation [6] und Evidenzgenerierung [7] spiegeln sich in den Ergebnissen verschiedener Übersichtsarbeiten wider, die auf eine eingeschränkte bis moderate Evidenz hinweisen [8, 9]. Dies bedeutet jedoch nicht, dass keine methodisch hochwertigen Studien durchgeführt wurden oder es keine Wirksamkeitsnachweise gibt [10]. Vielmehr sind bei der Vielzahl an Evaluationsstudien zu DI die methodischen Limitationen, wie meist geringe Fallzahlen, fast immer fehlende Verblindung, kurze Studiendauer sowie fehlende a priori veröffentlichte Studienprotokolle allgegenwärtig [11].

Vor dem Hintergrund dieser Herausforderungen wird in diesem zweiteiligen Diskussionspapier aus Perspektive der Arbeitsgruppe Digital Health des Deutschen Netzwerks Versorgungsforschung (DNVF) erörtert, welche theoretischen, praktischen und methodischen Herausforderungen mit der Evaluation von DI einhergehen. Es werden Lösungsansätze und zukünftige Forschungsbedarfe zur Adressierung dieser Herausforderungen aufgezeigt.

In diesem ersten Teil werden die Merkmale und Anwendungsgebiete, die Besonderheiten bei der Entwicklung und der Einbindung von Nutzer*innen sowie die Besonderheiten bei der Evaluation von DI in den Mittelpunkt gestellt.

Eigenschaften und Anwendungsfelder digitaler Interventionen

Bislang besteht eine begriffliche Unklarheit in Bezug auf gesundheitsbezogene DI [1]. Dies hängt mit den vielfältigen beteiligten Disziplinen, Akteur*innen und Nutzer*innen und methodischen Zugangswegen [2] sowie mit der Dynamik der Digitalisierung im Gesundheitswesen zusammen [12]. Es bestehen vielfältige Begrifflichkeiten (z. B. E-Health, Gesundheitstelematik), deren Bedeutung sich im zeitlichen Verlauf zum Teil verändert hat oder aber Begriffe synonym verwendet werden. In Konsequenz lassen sich unter *Digital Health* zunächst alle Informations- und Kommunikationstechnologien (IKT) aus dem Gesundheitsbereich zusammenfassen [4]. Eine Abgrenzung ist jedoch mit Blick auf spezifische Anwendungsbereiche, -formate oder -ziele möglich [13, 14].

Die WHO differenziert den Einsatz und Nutzen von DI für verschiedene Gruppen: 1) Nutzer*innen (Klient*innen oder Patient*innen) entsprechender Anwendungen, 2) Professionelle aus dem Gesundheitswesen, 3) mit Gesundheitsschutz, Prävention und Gesundheitsmanagement beauftragte Abteilungen bzw. Organisationen (z. B. Öffentlicher Gesundheitsdienst) sowie 4) Anbieter*innen bzw. Institutionen zur Sammlung, Auswertung und Darstellung gesundheitsbezogener Daten [2].

DI mit spezifischer Ausrichtung auf medizinische Versorgung sind der Telemedizin zuzuordnen. Diese wiederum kann hinsichtlich des medizinischen Fachgebiets (z. B. Telekardiologie, Telerradiologie) oder der genutzten Form der Anwendung (z. B. Telekonsultation, Telemonitoring) differenziert werden [15]. Telemedizin nutzt (1) IKT, (2) um eine Distanz zu überbrücken und (3) durch Leistungserbringende bei Patient*innen eine Versorgung (oder Schulung) durchzuführen [14]. Anwendungen im Bereich der psychiatrischen und psychotherapeutischen Versorgung werden mittlerweile auch als Ecological Momentary Interventions (EMIs) [16], Just in Time Adaptive Interventions (JITAs) [17] oder Internetgestützte und Mobilbasierte Interventionen (IMI) [18] bezeichnet. Weitere Bereiche stellen DI dar, die im Kontext von Prävention und Gesundheitsförderung eingesetzt werden [19]. Unter dem Einsatz mobiler gesundheitsbezogener Technologien (mHealth) ermöglichen Apps oder Sensoren im Smartphone das Bewegungs- oder Ernährungsverhalten kontinuierlich zu messen und zu beobachten. Diese Angebote haben in den letzten Jahren durch die Entwicklung von Gesundheits- bzw. Lifestyle-Apps und durch zertifizierte Medizin-Apps an Bedeutung zugenommen [4].

Mit dem Digitale-Versorgung-Gesetz besteht die Möglichkeit, digitale Gesundheitsanwendungen (DiGA) auf Rezept zu verordnen. Bei DiGAs handelt es sich um Medizinprodukte (MP) niedriger Risikoklasse, deren Hauptfunktion wesentlich auf digitalen Technologien beruht (§ 33a Abs. 1 SGB V). Diese durchlaufen einen Begutachtungs- und Erprobungsprozess unter Aufsicht des Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM), welcher als zügiger Fast-Track konzipiert ist.

Zusätzlich zu der Heterogenität von DI ist auch zu berücksichtigen, dass diese vielfach aus mehreren (teilweise interdependenten) Einzelkomponenten bestehen [20] und nicht solitär, sondern in Kombination mit weiteren (nicht-digitalen) Komponenten eingesetzt werden. Da viele DI zeit- und ortsunabhängig genutzt werden können [15] entfalten sie ihre Wirkung durch die Integration in komplexe soziale Systeme und Lebenswelten [21, 22]. Somit sind DI selbst komplexe Interventionen oder zumindest in ebensolche eingebettet [19]. Dies hat Einfluss auf die Planung, Durchführung, Analyse und Interpretation von Evaluationen. [19–20]

Entwicklung digitaler Interventionen

Die Entwicklung von DI ist komplex und planungsintensiv, insbesondere wenn ein Übergang in die Regelversorgung angestrebt wird. Hierbei besteht ein Spannungsfeld zwischen den Bedarfen und Wünschen zukünftiger Nutzer*innengruppen und dem technisch Möglichen [23]. Diesen Herausforderungen kann mit einem iterativen, nutzer*innenzentrierten und partizipativen Entwicklungsprozess begegnet werden [25]. Dieser orientiert sich an der ISO 9241–210:2019 und ist in vier Phasen untergliedert [25]. Mehrheitlich wird dieser Prozess mit dem Begriff User-Centered-Design bezeichnet, z.T. auch als Stakeholder-Centred Design.

- 1. Kontextanalyse:** Zu Beginn erfolgt die Analyse des Versorgungs-Kontextes einer DI und wer die Nutzer*innengruppen mit ihren objektiven Bedarfen und subjektiven Bedürfnissen sind. Die multidisziplinäre Kontextanalyse ist essentiell, da sich Strukturen und Prozesse durch die Implementierung von DI stark verändern können [24]. In der Entwicklungsphase kommen vor allem qualitative (z. B. Fokusgruppen) und quantitative Erhebungsmethoden (z. B. Fragebogenerhebungen) sowie Feldbeobachtungen zum Einsatz [25]. Aber auch die Auswertung von Routinedaten kann wertvolle Hinweise liefern. Zu einer umfassenden Kontextanalyse gehört eine Prüfung, welche DI oder Interventionskomponenten bereits existieren. Hierdurch werden Doppelentwicklungen vermieden und es wird an bereits existierende Entwicklungen angeknüpft.
- 2. Anforderungsanalyse:** Im zweiten Schritt werden die Anforderungen der Nutzer*innen an die DI definiert. Hierzu kommen u. a. Methoden wie sog. Personas (Nutzer*innenmodelle, die die Zielgruppe beschreiben) oder use cases (Beschreibung des Verhaltens eines Systems aus Anwender*innensicht) zum Einsatz [26].
- 3. Gestaltung:** Erst im Anschluss an die Analysen erfolgt der eigentliche Designprozess, u. a. mit Mock-ups (Anschauungsmodellen) und Card Sorting (Verfahren zur Entwicklung von Navigations- und Menüstrukturen), bis zum Prototypen [27].
- 4. Evaluation:** Nachdem Prototypen erstellt wurden, erfolgt eine erste formative Evaluation durch die potentiellen Nutzer*innen. Erfasst werden hier unter anderem die Effektivität und Effizienz sowie die allgemeine Zufriedenheit [24]. Hier kommen z. B. Usability-Tests (Test der Gebrauchstauglichkeit), Cognitive Walkthroughs (Evaluation eines Systems/Produkts aus Nutzer*innensicht durch Expert*innen) oder Thinking Aloud-Techniken (Methode zur Erfassung kognitiver Vorgänge bei der Bearbeitung von Testaufgaben) zum Einsatz [28]. Mit den Informationen aus der Evaluation wird die DI modifiziert und optimiert, anschließend erfolgt eine erneute Iteration der Evaluation (► **Abb. 1**) [25].

Neben dieser vorrangig produktzentrierten Evaluation sollte bei DI eine versorgungszentrierte Evaluation erfolgen. Hierbei sollten sowohl Prozess- als auch Ergebnisparameter unter Berücksichtigung medizinischer, technischer, sozialer und ökonomischer Aspekte einfließen [29]. Die Wahl der hierfür passenden Outcomes und Methoden hängt vom Entwicklungsstand/Reifegrad der Intervention ab [29, 30]. Beispielsweise sind in frühen Phasen der Entwicklung von DI eher akzeptanzorientierte Outcomes relevant, die oftmals mit qualitativen Methoden ermittelt werden. Hingegen sind in späteren Phasen eher Outcomes relevant, mit denen die Effektivität der Wirkung durch quantitative Methoden ermittelt werden (eine Vertiefung findet sich im nachfolgenden Absatz sowie in Teil II des Diskussionsartikels). Bei der Wahl der Methode sollte das der Risikostufe angemessene höchstmögliche Evidenzlevel gewählt werden [29, 31].

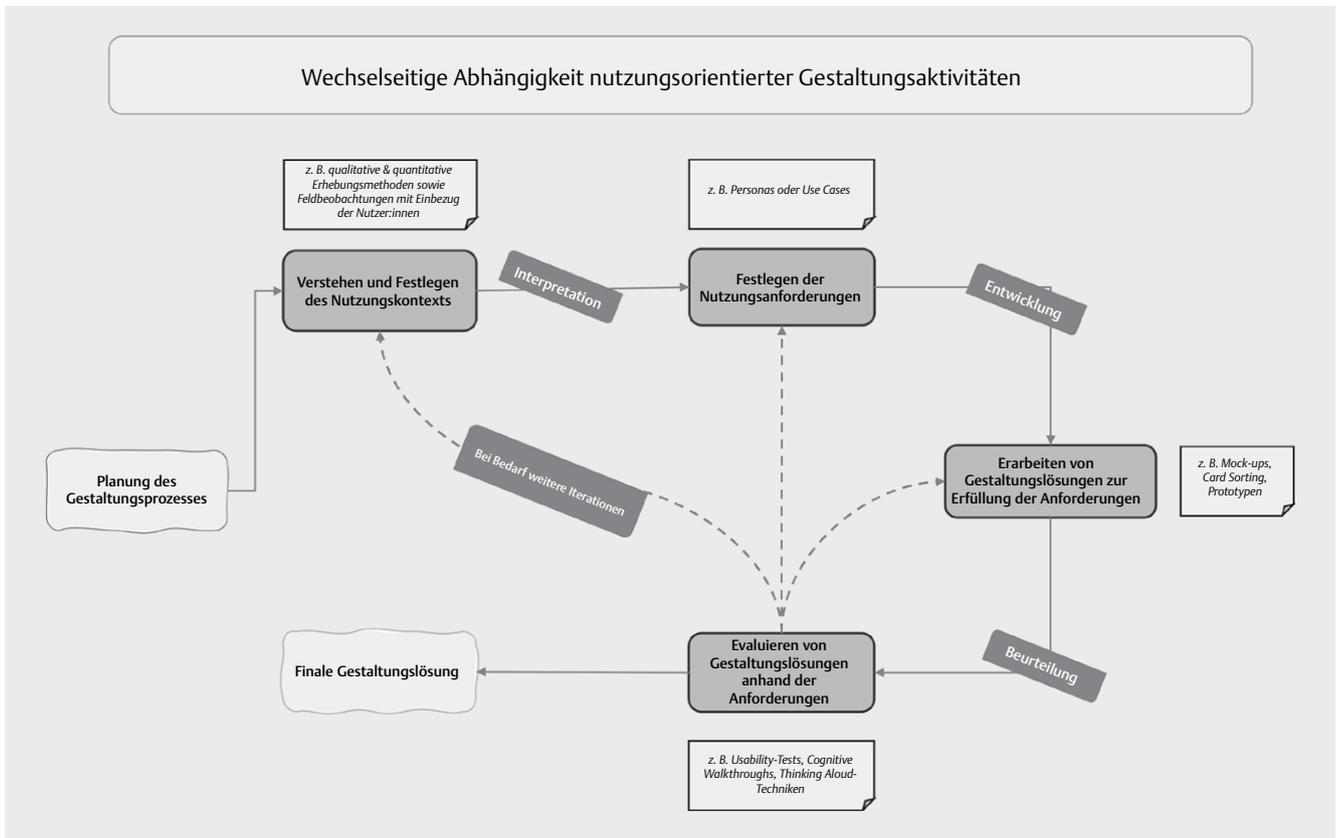
Ein Einbezug potentieller Nutzer*innengruppen sowie weiterer relevanter Akteur*innen, Interessengruppen und Entscheidungsträger von Anfang an ist eine wesentliche Voraussetzung für eine hohe Akzeptanz und damit die Wirksamkeit und nachhaltige Entwicklung und Implementierung von DI. Haupteinflussfaktoren der Akzeptanz sind entsprechend der etablierten Technologieakzeptanzmodelle, wie etwa der Unified Theory of Acceptance and Use of Technology oder Technology Acceptance Model, die empfundene Einfachheit der Nutzung sowie die empfundene Nützlichkeit/Zusatznutzen [32]. Ein Blick auf die aktuelle Studienlage zeigt jedoch, dass ein solches Vorgehen bei der Entwicklung digitaler Interventionen noch nicht die Regel ist. Ein Review von Maramba et al. [33] ergab, dass Usability-Studien nur für einen Teil der verfügbaren DI veröffentlicht sind. Zudem wurden bei der Mehrheit der eingeschlossenen Studien lediglich ein oder zwei Methoden im Rahmen der Evaluation eingesetzt, und zwar vorrangig quantitative Erhebungsmethoden. Dies restringiert eine umfassende Beurteilung und Berücksichtigung der Nutzer*innenanforderungen. Überdies fand bei der Mehrheit der Studien lediglich ein Iterationszyklus statt.

Ergänzend hierzu sollte bei der Entwicklung von DI von Beginn an eine potentielle Zertifizierung als MP und ein Übergang in die Regelversorgung berücksichtigt werden (siehe hierzu Teil II des Diskussionsartikels).

Ergänzend hierzu sollte bei der Entwicklung von DI von Beginn an eine potentielle Zertifizierung als MP und ein Übergang in die Regelversorgung berücksichtigt werden (siehe hierzu Teil II des Diskussionsartikels).

Evaluation digitaler Interventionen

Der Goldstandard für den Nutznachweis von Interventionen sind prospektive randomisiert kontrollierte Studien (RCT) [34]. Sie stoßen aufgrund der genuinen Merkmale von DI teilweise an ihre Grenzen [35]. Die Durchführung von RCTs setzt das Vorhandensein einer stabilen, inhaltlich ausgereiften Intervention [36] sowie ein hohes Maß an Standardisierung externer Faktoren bzw. deren Adjustierung gegenüber Verzerrungen voraus [34]. Es sollte eine begrün-



► **Abb. 1** Wechselseitige Abhängigkeit nutzungsorientierter Gestaltungsaktivitäten (Eigene Darstellung nach [24]).

dete Annahme bestehen, dass die Intervention mit einer hohen Wiedergabebetreue (delivery) durchgeführt werden kann und dass ein bedeutsamer und theoretisch begründbarer klinischer Effekt zu erwarten ist [6]. Häufig werden Evaluationsstudien zu DI nicht als RCT durchgeführt [7, 37], sondern lediglich im Pre-Post-Design [35]. Darüber hinaus werden die in systematischen Übersichtsarbeiten eingeschlossenen RCTs oftmals mit einer geringen bis moderaten Qualität bewertet [7, 38]. Faktoren, welche die Validität von RCT zu DI zusätzlich einschränken, sind kurze Beobachtungs- bzw. Interventionszeiträume [39] und hohe Drop-out-Raten [35, 40]. Gründe für letzteres sind u. a., dass die Zugehörigkeit zu einer Kontrollgruppe eher als bei anderen Interventionen auffällt und insbesondere redundante Inhalte auf Dauer frustrierend wirken können [41]. Die Durchführung von RCTs wird zusätzlich dadurch erschwert, dass eine Verblindung für Nutzer*innen und Leistungserbringer*innen, schwierig zu realisieren ist [42]. Es besteht jedoch die Möglichkeit, mit reduzierten Varianten der DI (oder auch einer funktionslosen Scheinapp) in der Kontrollgruppe zu operieren [43]. Die Nutzung einer Scheinapp reduziert hierbei zudem den potentiellen Einfluss des digitalen Placebo-Effekts und die Dropout-Raten, da beide Gruppen eine digitale Intervention erhalten [44].

In Bezug auf die Charakteristika von RCT im Kontext DI wird kritisiert, dass diese mit den besonderen Merkmalen von DI schwer vereinbar sind [45]. Adaptive Studiendesigns, die ein größeres Maß an Variabilität und Individualisierung ermöglichen und der Entwicklungsdynamik gerecht werden, gelangen zunehmend in die Dis-

kussion [3, 46]. Ein Merkmal adaptiver Studiendesigns ist, dass während der Durchführung Interimsanalysen durchgeführt werden, die zu einer (a priori geplanten) Adaption führen. Anpassungen, die nach einer solchen Interimsanalyse vorgenommen werden können, sind die Adjustierung der Stichprobengröße oder des Zuweisungsverhältnisses zwischen Studienarmen [46]. In diesem Kontext sind auch sog. Drop-the-Loser-Designs zu nennen, bei denen während der Interimsanalyse nachweislich ineffektive Interventionskomponenten aus weiteren Analysen ausgeschlossen werden [47]. Werden digitale Interventionen in der Regelversorgung, etwa in mehreren Kliniken, getestet, so ist auch eine Cluster-Randomisierung der Kliniken statt individueller Randomisierung denkbar, welche zudem systematischen Bias aufgrund der Clusterzugehörigkeit ausschließt [48]. Die Frustration darüber, in der Kontrollgruppe gelandet zu sein, wird in Stepped-Wegde-Designs vermieden, in dem Proband:innen zu einem definierten Zeitpunkt von der Kontroll- in die Interventionsgruppe wechseln können [49]. Auch wenn hierbei eine Verblindung nicht durchzuhalten ist, können Stepped-Wegde-Designs dazu beitragen Dropout-Raten zu minimieren [50]. Eine weitere Alternative bilden Studiendesigns, welche eine historische Kontrollgruppe, z. B. auf Basis der Abrechnungsdaten von Krankenkassen oder Daten aus krankheitsrelatierten Registern, bilden [51].

Weitere alternative Studiendesigns, wie etwa „Continuous Evaluation of Evolving Behavioral Intervention Technologies“ (CEEBIT), „Multiphase Optimization Strategy“ (MOST), „Sequential Multiple Assignment Randomized Trial“ (SMART) und „Micro-Randomized Trial“ (MRT) werden ebenso diskutiert [45]. Es besteht die Ansicht,

dass diese jedoch als nicht hinreichend für Wirksamkeits- und Nutzenbewertungen von DI angesehen werden, da sie nicht in der Lage sind einen der RCT vergleichbaren Evidenzgrad zu generieren [45]. Neben alternativen Studiendesigns ist eine stärkere Einbindung qualitativer Methoden in Parallelgruppen-Designs nützlich. Fokusgruppen vor und nach dem Abschluss von Interventionen könnten sich eignen, qualitative Veränderungen des medizinischen Nutzens (u. a. Gesundheitszustand und Lebensqualität) oder der Nutzenerwartung, etwa in Form von Patient-Reported Outcome Measure (PROMs) und Patient Reported Experience Measures (PREMs), sowie hemmende und fördernde Faktoren der DI aus Sicht der Nutzer*innen abzubilden [52]. Dies kann auch im Rahmen von Prozessevaluationen und mit Hilfe von Mixed-Methods-Designs geschehen.

Zudem sollten die Qualität und Wirkung von DI auch nach der Entwicklung und Implementation betrachtet werden, so wie es bei MPs gefordert wird (Post-Market-Surveillance) [3, 53]. Jedoch ist es auch für DI, die nicht als MP vorgesehen oder bereits zertifiziert sind, sinnvoll ein Monitoring der Wirkung und Qualität durchzuführen.

Obwohl RCTs der Goldstandard für die Wirksamkeitsprüfung von Interventionen im Gesundheitswesen sind, scheint bei der Vielzahl eine systematische Nutzenbewertung aller DI durch RCTs kaum praktikabel. Entscheidend sind die intendierten Funktionalitäten und das damit einhergehende Verhältnis von Nutzen und Schaden [54]. Das National Institute for Health and Care Excellence (NICE) hat hierzu das Evidence Standards Framework for Digital Health Technologies veröffentlicht (NICE-Framework), in dem DI nach Schaden- und Nutzenverhältnis in drei Evidenzstufen eingeordnet werden [31].

Zusammenfassend lässt sich derzeit festhalten, dass RCTs noch immer als Goldstandard für den Nutznachweis gelten. Für alternative Studiendesigns gilt, dass weitere Erfahrungen gesammelt werden müssen und deren Anwendung im Kontext von DI weiterzuentwickeln sind. Eine Erwägung alternativer Studiendesigns wird im Gutachten „Digitalisierung für Gesundheit“ des Sachverständigenrates zur Begutachtung der Entwicklung im Gesundheitswesen empfohlen und entsprechende Designs beschrieben [3]. Da adaptive Studiendesigns bis heute nur selten genutzt werden, ist zukünftig ebenso zu eruieren, wie sie zum Einsatz kommen können, ohne Abstufungen des Evidenzlevels in Kauf zu nehmen [55]. Das NICE-Framework kann für den deutschen Kontext ein Good-Practice-Beispiel darstellen [31, 42].

Schlussbetrachtung

Ziel dieses ersten Teils des Diskussionsartikels war es, aus Sicht der Versorgungsforschung zu diskutieren, welche Herausforderungen bei der Entwicklung und Evaluation von DI aufgrund deren spezifischen Eigenschaften existieren sowie Lösungsansätze und zukünftige Forschungsbedarfe aufzuzeigen.

Versorgungsforschung hat als Ziel, die Forschung zu Versorgungsstrukturen und Versorgungsprozessen in tatsächlichen und alltäglichen Abläufen zu adressieren [56]. Daraus folgt, dass eine kontinuierliche Auseinandersetzung mit der Evaluation von DI unabdingbar ist. Insbesondere die Komplexität sowie die spezifischen

Möglichkeiten und Herausforderungen von DI stellen besondere Anforderungen an die Evaluation von DI. Betrachtet man jedoch die Empfehlungen zur Entwicklung und Evaluation komplexer Interventionen, so adressieren diese die diskutierten besonderen Eigenschaften von DI noch nicht hinreichend genug [6, 57].

In dem zweiten Teil des Diskussionsartikels werden weiterer Themenfelder diskutiert, die von Relevanz bei der Evaluation von DI sind. Hierzu zählen die Auswahl von Outcomes bei der Evaluation, dem Reporting von Studienergebnissen, der Synthese der Evidenz sowie der Implementierung von DI in den Versorgungsalltag.

Interessenkonflikt

Die Autorinnen/Autoren geben an, dass kein Interessenkonflikt besteht.

Literatur

- [1] Zeeb H, Pigeot I, Schüz B. Digital Public Health – ein Überblick. Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz 2020; 63: 137–144
- [2] WHO. Classification of Digital Health Interventions v1.0A shared language to describe the uses of digital technology for health [Internet]. 2018. Available from: <http://apps.who.int/iris/bitstream/handle/10665/260480/WHO-RHR-18.06-eng.pdf;jsessionid=E06A96F29EE14D51614B4C25590A04D7?sequence=1>
- [3] SVR-Gesundheit. Digitalisierung für Gesundheit Ziele und Rahmenbedingungen eines dynamisch lernenden Gesundheitssystems [Internet]. 2021. Available from: https://www.svr-gesundheit.de/fileadmin/Gutachten/Gutachten_2021/SVR_Gutachten_2021_online_.pdf
- [4] Kramer U, Borges U, Fischer F et al. DNVF-Memorandum – Gesundheits- und Medizin-Apps (GuMAs). Das Gesundheitswesen 2019; 81: e154–e170
- [5] Vollmar H, Kramer U, Müller H et al. Digitale Gesundheitsanwendungen – Rahmenbedingungen zur Nutzung in Versorgung, Strukturentwicklung und Wissenschaft – Positionspapier der AG Digital Health des DNVF. Das Gesundheitswesen 2017; 79: 1080–1092
- [6] Murray E, Hekler EB, Andersson G et al. Evaluating Digital Health Interventions Key Questions and Approaches. Am J Prev Med 2016; 51: 843–851
- [7] Byambasuren O, Sanders S, Beller E et al. Prescribable mHealth apps identified from an overview of systematic reviews. Npj Digital Medicine 2018; 1: 12
- [8] Kolasa K, Kozinski G. How to Value Digital Health Interventions? A Systematic Literature Review. Int J Environ Res Pu 2020; 17: 2119
- [9] Tokgöz P, Hrynyschyn R, Hafner J et al. Digital health interventions in prevention, relapse, and therapy of mild and moderate depression: scoping review. JMIR Mental Health 2021; 8(4): e26268. DOI: 10.2196/26268
- [10] Mathews S, D’Arcey J, Tokgöz P et al. Digital Health Interventions in Prevention, Relapse, and Therapy of Mild and Moderate Depression: Scoping Review. Jmir Ment Heal 2021; 8: e26268
- [11] Timpel P, Oswald S, Schwarz PEH et al. Mapping the Evidence on the Effectiveness of Telemedicine Interventions in Diabetes, Dyslipidemia, and Hypertension: An Umbrella Review of Systematic Reviews and Meta-Analyses. J Med Internet Res 2020; 22: e16791

- [12] Shaw T, McGregor D, Brunner M et al. What is eHealth (6)? Development of a Conceptual Model for eHealth: Qualitative Study with Key Informants. *J Med Internet Res* 2017; 19: e324
- [13] WHO. National eHealth Strategy Toolkit Overview [Internet]. 2012. Available from: <https://www.who.int/ehealth/publications/overview.pdf>
- [14] Otto L, Harst L, Schlieter H et al. Towards a Unified Understanding of eHealth and Related Terms – Proposal of a Consolidated Terminological Basis. *Proc 11th Int Jt Conf Biomed Eng Syst Technologies* 2018; 533–539
- [15] Fischer F, Aust V, Krämer A. eHealth in Deutschland, Anforderungen und Potenziale innovativer Versorgungsstrukturen 2016; 3–23
- [16] Reininghaus U. Ambulatorische Interventionen in der Psychiatrie: das Momentum für Veränderung im alltäglichen sozialen Kontext. *Psychiatr Prax* 2018; 45: 59–61
- [17] Nahum-Shani I, Smith SN, Spring BJ et al. Just-in-Time Adaptive Interventions (JITAs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Ann Behav Med* 2016; 52: 1–17
- [18] Lin J, Baumeister H. Internet- und Mobilebasierte Interventionen in der Psychotherapie. *Public Heal Forum* 2015; 23: 176–179
- [19] Fischer F. Digitale Interventionen in Prävention und Gesundheitsförderung: Welche Form der Evidenz haben wir und welche wird benötigt? *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz* 2020; 1–7
- [20] Mühlhauser I, Lenz M, Meyer G. Bewertung von komplexen Interventionen: Eine methodische Herausforderung. 2012; 109 (1–2): A-22. Available from: <https://www.aerzteblatt.de/int/article.asp?id=118129>
- [21] Robert-Koch-Institut, Lebensmittelsicherheit BL für G und. Evaluation komplexer Interventionsprogramme in der Prävention: Lernende Systeme, lehrreiche Systeme? [Internet]. 2012. Available from: https://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsB/evaluation_praevention.pdf?__blob=publicationFile
- [22] Koppe G, Guloksuz S, Reininghaus U et al. Recurrent Neural Networks in Mobile Sampling and Intervention. *Schizophrenia Bull* 2018; 45: 272–276
- [23] Dockweiler C, Razum O. Digitalisierte Gesundheit: neue Herausforderungen für Public Health. *Das Gesundheitswesen* 2015; 78: 5–7
- [24] ISO. ISO 9241-210:2019(en) Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems [Internet]. 2019. Available from: <https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en>
- [25] Diekmann A. Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen 2020
- [26] Goodman E, Kuniavsky M, Moed A. Observing the User Experience 2012
- [27] Sherwin K. Card Sorting: Uncover Users' Mental Models for Better Information Architecture. 2018; Available from: <https://www.nngroup.com/articles/card-sorting-definition/>
- [28] Baxter K, Courage C, Caine K. Understanding your Users. *A Practical Guide to User Research Methods* 2015
- [29] Arnold K, Scheibe M, Müller O et al. Principles for the evaluation of telemedicine applications: Results of a systematic review and consensus process. *Zeitschrift Für Evidenz Fortbildung Und Qual Im Gesundheitswesen*. 2016; 117: 9–19
- [30] Kowatsch T, Otto L, Harperink S et al. A design and evaluation framework for digital health interventions. *Information Technology* 2020; 61: 253–263
- [31] NICE. Evidence standards framework for digital health technologies. 2019. <https://www.nice.org.uk/corporate/ecd7>
- [32] Venkatesh V, Morris M, Davis G et al. User acceptance of information technology: Toward a unified view. *MIS Quarterly* 2003; 27: 425–478
- [33] Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: A scoping review. *Int J Med Inform* 2019; 126: 95–104
- [34] Peiris D, Miranda JJ, Mohr DC. Going beyond killer apps: building a better mHealth evidence base. *Bmj Global Heal* 2018; 3: e000676
- [35] Guo C, Ashrafian H, Ghafur S et al. Challenges for the evaluation of digital health solutions – A call for innovative evidence generation approaches. *Npj Digital Medicine* 2020; 3: 110
- [36] Henderson M, Donnachie C. Improving eHealth intervention development and quality of evaluations. *The Lancet Digital Health* 2019; 1: e194–e195
- [37] Pawloski PA, Brooks GA, Nielsen ME et al. A Systematic Review of Clinical Decision Support Systems for Clinical Oncology Practice. *J Natl Compr Canc Ne* 2019; 17: 331–338
- [38] Champion KE, Parmenter B, McGowan C et al. Effectiveness of school-based eHealth interventions to prevent multiple lifestyle risk behaviours among adolescents: a systematic review and meta-analysis. *Lancet Digital Heal* 2019; 1: e206–e221
- [39] Vilardaga R, Casellas-Pujol E, McClernon JF et al. Mobile Applications for the Treatment of Tobacco Use and Dependence. *Curr Addict Reports* 2019; 6: 86–97
- [40] Linardon J, Fuller-Tyszkiewicz M. Attrition and Adherence in Smartphone-Delivered Interventions for Mental Health Problems: A Systematic and Meta-Analytic Review. *J Consult Clin Psych* 2019; 88: 1–13
- [41] Reinhardt G, Schwarz PE, Harst L. Non-use of telemedicine: A scoping review. *Health Inform J*. 2021; 27: 14604582211043148
- [42] Neugebauer EAM, Rath A, Antoine S-L et al. Specific barriers to the conduct of randomised clinical trials on medical devices. *Trials*. 2017; 18: 427
- [43] Timpel P, Cesena FHY, da Silva Costa C et al. Efficacy of gamification-based smartphone application for weight loss in overweight and obese adolescents: study protocol for a phase II randomized controlled trial. *Ther Adv Endocrinol Metabolism* 2018; 9: 167–176
- [44] Stalujanis E, Neufeld J, Stalder MG et al. Induction of Efficacy Expectancies in an Ambulatory Smartphone-Based Digital Placebo Mental Health Intervention: Randomized Controlled Trial. *Jmir Mhealth Uhealth* 2021; 9: e20329
- [45] Gensorowsky D, Lampe D, Hasemann L et al. „Alternative Studiendesigns“ zur Bewertung digitaler Gesundheitsanwendungen – eine echte Alternative? *Zeitschrift Für Evidenz Fortbildung Und Qual Im Gesundheitswesen* 2021; 161: 33–41
- [46] Pallmann P, Bedding AW, Choodari-Oskooei B et al. Adaptive designs in clinical trials: why use them, and how to run and report them. *Bmc Med* 2018; 16: 29
- [47] Chen YHJ, DeMets DL, Lan KKG. Some drop-the-loser designs for monitoring multiple doses. *Stat Med* 2010; 29: 1793–1807
- [48] Puffer S, Torgerson DJ, Watson J. Cluster randomized controlled trials. *J Eval Clin Pract* 2005; 11: 479–483
- [49] Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007; 28: 182–191
- [50] Beard E, Lewis JJ, Copas A et al. Stepped wedge randomised controlled trials: systematic review of studies published between 2010 and 2014. *Trials* 2015; 16: 353
- [51] (BfArM) B für A und M. DiGA-Verzeichnis [Internet]. 2021. Available from: <https://diga.bfarm.de/de>
- [52] Ekeland AG, Bowes A, Flottorp S. Methodologies for assessing telemedicine: A systematic review of reviews. *Int J Med Inform* 2012; 81: 1–11

- [53] Wilson K, Bell C, Wilson L et al. Agile research to complement agile development: a proposal for an mHealth research lifecycle. *Npj Digital Medicine* 2018; 1: 46
- [54] Angelescu K, Sauerland S. Mobile Gesundheitsanwendungen: Welche Evidenz ist n-tig? *Dtsch Arztebl International* [Internet] 2019; 116:A-1057 Available from: <https://www.aerzteblatt.de/int/article.asp?id=207821>
- [55] Dick S, O'Connor Y, Heavin C. Approaches to Mobile Health Evaluation: A Comparative Study A mHealth Evaluation Comparison Study. *Inform Syst Manage* 2019; 37: 1–18
- [56] Philipp S-P. Digitalisierung der Versorgungsforschung – Versorgungsforschung zur Digitalisierung. In Berlin, Heidelberg. Springer; Berlin Heidelberg: 2019: p 175–184
- [57] Moore GF, Evans RE, Hawkins J et al. From complex social interventions to interventions in complex social systems: Future directions and unresolved questions for intervention development and evaluation. *Evaluation-us* 2019; 25: 23–45