

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/351282491>

# Bewerten, aber wie? – Kriterien für die Evaluation und das Reporting von Studien zu Gesundheits- und Medizin-Apps

Research · March 2021

DOI: 10.24945/MVF.02.21.1866-0533.2300

CITATION

1

READS

359

4 authors, including:



**Sven Kernebeck**  
FH Münster

69 PUBLICATIONS 189 CITATIONS

[SEE PROFILE](#)



**Marcus Redaelli**  
Universität Witten/Herdecke

15 PUBLICATIONS 44 CITATIONS

[SEE PROFILE](#)



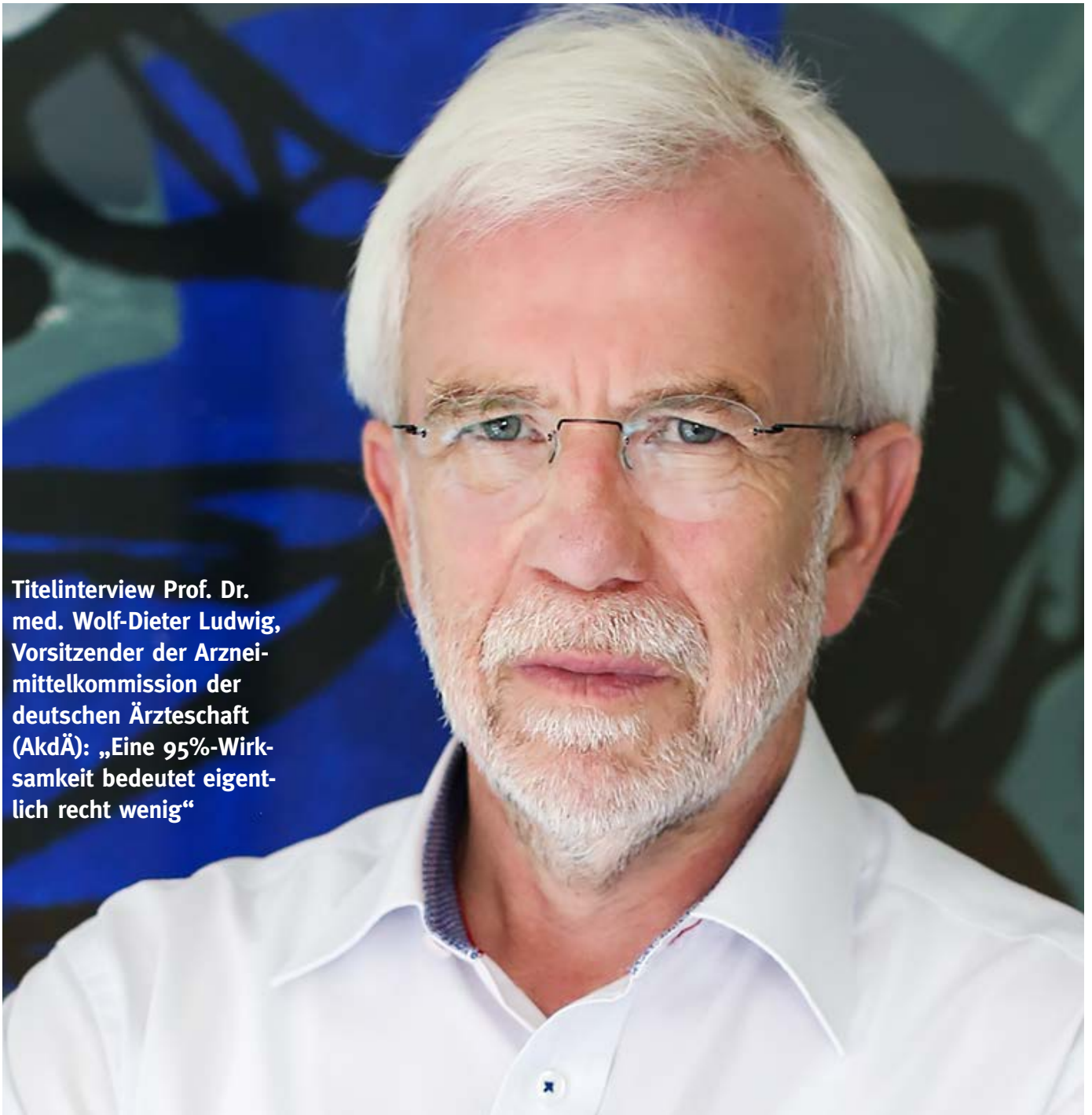
**Horst Christian Vollmar**  
Ruhr-Universität Bochum

237 PUBLICATIONS 1,561 CITATIONS

[SEE PROFILE](#)

# monitor **VERSORGUNG FORSCHUNG**

Fachzeitschrift zu Realität, Qualität und Innovation der Gesundheitsversorgung



Titelinterview Prof. Dr. med. Wolf-Dieter Ludwig, Vorsitzender der Arzneimittelkommission der deutschen Ärzteschaft (AkdÄ): „Eine 95%-Wirksamkeit bedeutet eigentlich recht wenig“

- „Klinisch-genomische Daten für Forschung und Versorgung“ (Westphalen/Janssen)
- „Gesundheitssystem neu denken“ (Piwernetz/Neugebauer)
- „Jetzt einmalige Chance für Covid-19-Forschung nutzen“ (Gothe/Kurepkat)

## Editorial

### Covid-19: Das zweite Jahr 4

Herausgeber Prof. Dr. Reinhold Roski

## Redaktion

### „95%-Wirksamkeit bedeutet eigentlich recht wenig“ 6

Titelinterview mit Prof. Dr. med. Wolf-Dieter Ludwig, Vorsitzender der AkdÄ

### Zero- ist nicht No-, aber fast Low-Covid 14

Strategieansätze zur Bekämpfung der Corona-Pandemie

### Ausgewählte Kollateralschäden von Covid-19 16

Die Fallzahlentwicklung von Herz- und Hirninfarkten wirft Fragen auf

### Das Corona-Paradoxon 18

2. Charité-Versorgungsforschungskongress

### Auf der Suche nach über 13 Milliarden Euro 20

Lehren aus der Pandemie für eine grundlegende Krankenhausreform

### Möglichkeiten/Grenzen der Covid-19-Modellierung 22

Eine Übersicht über Covid-19-Modellierungen und deren Modellierer

### Einmalige Chance für Covid-19-Forschung nutzen 28

Warum wir mehr ambulante Daten in Covid-19-Registern brauchen

### UPD geht aktiv auf die Versorgungsforschung zu 32

„Modell 2023“ zur Weiterentwicklung der Unabhängigen Patientenberatung

### „Klinisch-genomische Daten“ 34

Im Interview: Dr. Benedikt Westphalen (CCC MünchenLMU) und Dr. Katja Janssen (Roche)

### Kunst der „Evaluation komplexer Interventionen“ 38

Teil 3: Online-Fachtagung der IGiB StimT

### „Gesundheitssystem neu denken“ 40

Kommentar von Dr. Dr. Klaus Piwernetz, Univ.-Prof. Dr. Dr. h.c. E.A.M. Neugebauer

### „Corona-Virus als Stresstest“ 44

Kommentar von Prof. Dr. rer. oec. Axel C. Mühlbacher, Prof. Dr. oec. publ. Peter Zweifel

### Vorschläge zur Priorisierung elektiver Eingriffe 48

1. Diskussions- und Arbeitspapier der Marburger Bund-Fraktion der Ärztekammer Berlin

### Priorisierung kinderchirurgischer Eingriffe 51

2. Diskussions- und Arbeitspapier der Marburger Bund-Fraktion der Ärztekammer Berlin

### „Deutlich bessere Sichtbarkeit für die Pflege 54

Serie (Teil 30): Abteilung Pflegewissenschaftliche Versorgungsforschung, Bremen

## Zahlen - Daten - Fakten

### MS: Therapiebestimmend ist die Komplexität 12

## Standards

Impressum 2 News 43, 47 Rezensionen 58

## Impressum

Monitor Versorgungsforschung – Fachzeitschrift zu Realität, Qualität und Innovation der Gesundheitsversorgung

**Monitor Versorgungsforschung**  
Fachzeitschrift zu Realität,  
Qualität und Innovation der  
Gesundheitsversorgung  
14. Jahrgang  
ISSN: 1866-0533 (Printversion)  
ISSN: 2509-8381 (eFirst)

**Herausgeber**  
Prof. Dr. Reinhold Roski, Berlin  
roski@m-vf.de

**Chefredaktion**  
Peter Stegmaier (verantwort. Redakt.)

Kölstr. 119, 53111 Bonn  
Tel +49 228 7638280-0  
Fax +49 228 7638280-1

stegmaier@m-vf.de  
**Redaktion**

Kerstin Müller  
mueller@m-vf.de  
Jutta Mutschler  
mutschler@m-vf.de  
Martin Klein (Freier Journalist)  
klein@m-vf.de

**Verlag**  
eRelation AG – Content in Health  
Vorstand: Peter Stegmaier  
Kölstr. 119, 53111 Bonn  
www.ereRelation.org  
mail@ereRelation.org

**Verlagsleitung**  
Peter Stegmaier

**Anzeigenleitung/Vertrieb/Abo**  
Anke Heiser (verantwortlich für  
den Anzeigenteil)

heiser@m-vf.de  
**Marketing:**  
Kölstr. 119, 53111 Bonn  
Tel +49 228 7638280-0  
Fax +49 228 7638280-1

**Abonnement**  
„Monitor Versorgungsforschung“  
erscheint sechsmal jährlich. Der  
Preis für ein Jahresabonnement  
beträgt 120 Euro. Jahresvorzugspreis  
für Studenten gegen Vorlage einer  
Immatrikulationsbescheinigung 90  
Euro. zzgl. MwSt. und Versandkosten:  
Inland 9,99 Euro; Ausland 54 Euro.  
Preisänderungen vorbehalten. Die  
Abonnementdauer beträgt ein Jahr.  
Das Abonnement verlängert sich

automatisch um ein weiteres Jahr,  
wenn es nicht spätestens sechs  
Wochen vor Ablauf des Bezugsjahres  
schriftlich gekündigt wird.

**Verpackung**  
Die Verpackung dieser Zeitschrift  
ist bei www.verpackungsregister.  
org (LUCID) registriert unter:  
DE3360908810552

**Layout**  
eRelation AG, Bonn  
**Druck**  
Kössinger AG & Co. KG  
Fruehaufstraße 21  
84069 Schierling

info@koessinger.de  
Tel +49-(0)9451-499140  
Fax +49-(0)9451-499101  
Printed in Germany

**Urheber- und Verlagsrecht**  
Die Zeitschrift und alle in ihr  
enthaltenen einzelnen Beiträge und  
Abbildungen sind urheberrechtlich  
geschützt. Mit Annahme des  
Manuskripts gehen das Recht zur  
Veröffentlichung sowie die Rechte zur  
Übersetzung, zur Vergabe von  
Nachdruckrechten, zur elektronischen  
Speicherung in Datenbanken, zur  
Herstellung von Sonderdrucken,  
Fotokopien und Mikrokopien an  
den Verlag über. Jede Verwertung

außerhalb der durch das  
Urheberrechtsgesetz  
festgelegten Grenzen ist  
ohne Zustimmung des  
Verlags unzulässig. In  
der unaufgeforderten Zusendung von  
Beiträgen und Informationen an den  
Verlag liegt das jederzeit widerrufliche  
Einverständnis, die zugesandten Bei-  
träge bzw. Informationen in Daten-  
banken einzustellen, die vom Verlag  
oder Dritten geführt werden.  
**Auflagenmeldung**  
Mitglied der Informationsgemeinschaft  
zur Feststellung der Verbreitung von  
Werbeträgern e.V. (IWV), Berlin. Ver-  
br. Auflage: 6.998 (IWV 4. Qu. 2020).



## WISSENSCHAFT

### Dr. phil. Stefan Peters / Prof. Dr. Sportwiss. Andrea Schaller / Prof. Dr. phil. habil. Christian Thiel / Prof. Dr. phil. Lars Gabrys / Dr. Sportwiss. Gisela Nellessen-Martens Körperliche Aktivität in der SARS-CoV-2-Pandemie: Selektiver Überblick und Perspektiven 63

Aus dem vorliegenden Überblick lassen sich erste Bedarfe und Perspektiven für die bewegungsbezogene Versorgungsforschung und Gesundheitspolitik ableiten.

### Sven Kernebeck MSc PH / Dr. Ursula Kramer / Dr. med. Marcus Redaelli / Prof. Dr. med. Horst Christian Vollmar MPH Bewerten, aber wie? – Kriterien für die Evaluation und das Reporting von Studien zu Gesundheits- und Medizin-Apps 69

Angesichts der in dieser Arbeit dargestellten Optionen von Standards in der Berichterstattung von digitalen Anwendungen konnte der vom BfArM vorgestellte Leitfaden um weitere Items erweitert werden, um Leistungserbringern und Anwendern mehr Sicherheit zu geben und gleichzeitig Abstriche im Qualitätsanspruch zu verhindern.

### Dr. rer. pol. Hannes Schlieter / Dr. rer. medic. Patrick Timpel / Dr. rer. pol. Lena Otto / Peggy Richter MSc / Dipl.-Ing. Bastian Wollschlaeger / Andreas Knapp MSc / Lorenz Harst MA Digitale Gesundheitsanwendungen – Forderungen für deren Entwicklung, Implementierung und begleitende Evaluation 76

Trotz ihres allgemein anerkannten Potenzials scheitern auch in Deutschland Digitale Gesundheitsanwendungen (DiGA) nach wie vor an der nachhaltigen Integration in die Regelversorgung. Im Folgenden werden, im Einklang mit den Anforderungen des neu eingerichteten BfArM-Fast-Tracks für die Aufnahme von DiGA in die Regelversorgung, mehrere konkrete Maßnahmen zur Standardisierung des Evaluations- und Implementierungsprozesses vorgeschlagen und anhand eines prototypischen Implementierungsprozesses visualisiert.

### Dr. rer. nat. Luis Möckel / Prof. Dr. phil. Yvonne Treusch Medikamentöse und nicht-medikamentöse Vergiftungen in Deutschland: eine Analyse von Krankenhausdiagnosedaten aus der GBE-Datenbank 81

Ziel war die Häufigkeit von Krankenhausdiagnosen medikamentöser (T36.\*–T50.\*) und nicht-medikamentöser Vergiftungen (T51.\*–T65.\*) in Deutschland über die Jahre 2008 bis 2017 zu untersuchen. Korrelationen zwischen der Abnahme der Vergiftungsfälle wurden mit dem Pro-Kopf-Einkommen, der Arbeitslosenquote, der Anzahl der Haushalte mit Wohngeld sowie der Schulbildung für die Jahre 2008 bis 2017 ermittelt.


Bitte beachten Sie die Sonderveröffentlichung des  
DNVF auf den Seiten 59-62

Dieser Ausgabe liegen in Teilaufgaben „Monitor Pflege“ und „Pharma Relations“ sowie das „Lotsenplakat“ des BMC bei

Hinweis: Obwohl in MVF generell die männliche Schreibweise  
verwendet wird, sind immer alle Geschlechter gemeint.

# Herausgeber-Beirat

Universitäten/Hochschulen

	Arzneimittelkommission der deutschen Ärzteschaft	Prof. Dr. Wolf-Dieter Ludwig	
	Hochschule Neubrandenburg University of Applied Sciences	Prof. Dr. Axel C. Mühlbacher	
	MEDIZINISCHE HOCHSCHULE BRANDENBURG	Univ.-Prof. Dr. Prof. h.c. Dr. h.c. Edmund A.M. Neugebauer	
	Universität zu Köln	Prof. Dr. Matthias Schrappe	
	Universität zu Köln	Prof. Dr. Stephanie Stock	

Institute/Stiftungen

	bhwfib Forschungsinstitut	Prof. Dr.-Ing. Thomas P. Zahn	
	BDI	Felix Esser	
	BertelsmannStiftung	Uwe Schwenk	

Akteure

	Gesundheitsforen	Roland Nagel	
	AKRONBÜNDNIS PATIENTENSICHERHEIT	Dr. Ilona Köster-Steinebach	
	AOK Nordost	Harald Möhlmann	
	AOK Baden-Württemberg	Johannes Bauernfeind	
	B.A.G SELBSTHILFE	Dr. Martin Danner	
	Boehringer Ingelheim	Dr. Marco Penske	
	BKK Dachverband	Franz Knieps	
	DAK Gesundheit	Andreas Storm	
	Deutsche RHEUMA-LIGA GEMEINSAM MEHR BEWEGEN	Dr. Rolf Ulrich Schlenker	
	DiHeSys Digital Health Systems	Prof. Dr. Christian Franken	
	DocMorris Medikamente allein sind nicht genug	Sebastian Mindt	
	HCSG Health Care Systems GmbH	Dr. Peter Leiter	
	IGiB StimMT	Lutz O. Freiberg	
	IGV RESEARCH	Prof. Dr. rer.pol. h.c. Herbert Rebscher	

	socium Forschungszentrum Lebenswelt und Sozialpolitik	Prof. Dr. Gerd Glaeske	
	TUM	Prof. Dr. Leonie Sundmacher	
	UNIVERSITÄT BAYREUTH	Univ.-Prof. Dr. rer. pol. Dr. sc. nat. (ETH) Klaus H. Nagels	
	Universitätsmedizin	Prof. Dr. Wolfgang Hoffmann, MPH	

	IGES	Prof. Dr. Bertram Häussler	
	inav	Prof. Dr. Volker Amelung	
	Zi ZENTRALINSTITUT FÜR DIE KASSENÄRZTLICHE VERSORGUNG IN DEUTSCHLAND	Dr. Dominik von Stillfried	
	INSIGHT HEALTH	Christian Luley	
	janssen PHARMACEUTICAL COMPANIES OF Johnson & Johnson	Dr. Dorothee Brakmann	
	KVB Kassenärztliche Vereinigung Bayern	Dipl. Kfm. Fabian Demmelhuber	
	KVBB Kassenärztliche Vereinigung Brandenburg	MUDr./CS Peter Noack	
	Lilly	Dr. Thomas M. Zimmermann	
	MEDICAL CONTACT	Prof. Dr. Stephan Burger	
	NOVARTIS	Dr. Andreas Kress	
	OptiMedis	Dr. Oliver Gröne	
	Pfizer	Friedhelm Leverkus	
	Robert Bosch Stiftung	Dr. Bernadette Klapper	
	Roche	Dr. David Traub	
	SANOFI	Dr. Stephanie Rosenfeld, MHBA	
	Vivantes	Prof. Dr. Dr. Alfred Holzgreve	
		Dr. Christopher Hermann	





**Prof. Dr.  
Reinhold Roski**

Wirtschaftskommunikation  
und Gesundheitskommunikation  
Herausgeber von „Monitor  
Versorgungsforschung“

## Covid-19: Das zweite Jahr

**MVF-Titelinterview mit Prof. Dr. med. Wolf-Dieter Ludwig, Vorsitzender der Arzneimittelkommission der deutschen Ärzteschaft** > 6 ff.

Was sagen uns die Ergebnisse der Zulassungsstudien für die Impfstoffe gegen Covid 19? „Leider eher wenig.“ So **Professor Wolf-Dieter Ludwig**, Vorsitzender der Arzneimittelkommission der deutschen Ärzteschaft, im Titelinterview. „Sicher ist, dass die Impfungen nutzen, ... . Unklar bleibt jedoch – da eine systematische Testung aller Probanden mittels PCR auf SARS-CoV-2 Infektion nicht erfolgte –, ob eventuell mit BioNTech/Pfizer geimpfte Probanden z. B. Fieber und/oder Muskelschmerzen seltener gemeldet haben, weil sie von ihnen als Nebenwirkungen der Impfung und nicht als Symptome von Covid-19 erachtet wurden.“

**Neu im Herausgeberbeirat: Professor Volker Amelung, inav** > S. 3

Wir begrüßen als neues Mitglied im Herausgeberbeirat: **Prof. Dr. oec. Volker Amelung**, Geschäftsführer, Gründer und Gesellschafter des inav – privates Institut für angewandte Versorgungsforschung. Herr Amelung führt damit die Arbeit von Hans-Holger Bleß im Beirat fort. Wir freuen uns sehr über seine Mitwirkung und bedanken uns herzlich für das Engagement von Herrn Bleß.

**Covid-19-Pandemie** > S. ab S. 14.

Zahlreiche weitere Beiträge beschäftigen sich mit der laufenden Covid-19-Pandemie. Z.B. der Report über die Strategieansätze Zero-Covid und No-Covid und der Überblick zu Covid-19-Modellierungen sowie über Covid-19-Register (**Gothe/Kurepkat**). Wichtig auch die beiden Kommentare von **Neugebauer/Piwernetz** und **Mühlbacher/Zweifel** sowie zwei Diskussions- und Arbeitspapiere der Marburger Bund-Fraktion der Ärztekammer Berlin.

> S. 28 ff.  
> S. 40 ff., 44 ff.  
> S. 48 ff., 51 ff.

**Interview mit Dr. Benedikt Westphalen, CCC München LMU, und Dr. Katja Janssen, Roche Pharma AG** > S. 34 ff.

**Dr. Benedikt Westphalen**, der ärztliche Leiter der Präzisionsonkologie am CCC München<sup>LMU</sup>, und **Dr. Katja Janssen**, Personalised Healthcare Implementation Lead der Roche Pharma AG, plädieren dafür, klinisch-genomische Daten besser zu erfassen und zu nutzen.

**MVF-Serie „Inside Versorgungsforschung“** > S. 54 ff.

Diesmal sind wir bei Frau **Prof. Dr. Karin Wolf-Ostermann** und ihren Mitarbeiterinnen in der Abteilung Pflegewissenschaftliche Versorgungsforschung am Institut für Public Health und Pflegeforschung der Universität Bremen.

### Wissenschaftliche Beiträge

**Peters u.a.** untersuchen die Veränderungen der körperlichen Aktivität der Bevölkerung während des Lockdowns aufgrund der Covid-19-Pandemie. Zumindest einige Bevölkerungsgruppen konnten ihre körperliche Gesamtaktivität aufrechterhalten. > S. 63 ff.

**Kernebeck u.a.** analysieren sieben Publikationen zu Framework, Phasen/Items, Methodik und Fokus zur Evaluierung von Gesundheits- und Medizin-Apps und ergänzen damit den vom BfArM vorgestellten Leitfaden. > S. 69 ff.

**Schlieter u.a.** schlagen konkrete Maßnahmen zur Standardisierung des Evaluations- und Implementierungsprozesses von Digitalen Gesundheitsanwendungen vor, um eine bessere Integration in die Regelversorgung zu erreichen. > S. 76 ff.

**Möckel und Treusch** analysieren die Krankenhausdiagnosen medikamentöser und nicht-medikamentöser Vergiftungen in Deutschland 2008-2017. Präventive Maßnahmen zu medikamentösen Vergiftungen erscheinen für die Altersgruppe der 15- bis 24-Jährigen, für Frauen sowie finanziell- und bildungsbenachteiligte Gruppen angebracht. > S. 81 ff.

Vor allem auch im zweiten Jahr der Pandemie: Bleiben Sie gesund!

Mit herzlichen Grüßen  
Ihr Professor Dr. Reinhold Roski

Sven Kernebeck MSc PH  
 Dr. Ursula Kramer  
 Dr. med. Marcus Redaelli  
 Prof. Dr. med. Horst Christian Vollmar MPH

## Bewerten, aber wie? – Kriterien für die Evaluation und das Reporting von Studien zu Gesundheits- und Medizin-Apps

Mit der Digitalisierung im Gesundheitswesen sind große Erwartungen an die Qualitätssteigerung der gesundheitlichen Versorgung geknüpft (Mathews et al., 2019), dies gilt besonders für Diagnostik, Prävention sowie für das Management chronischer Krankheiten (Gordon, Landman, Zhang & Bates, 2020). Gesundheits- und Medizin-Apps (GuMA) werden aufgrund der ubiquitären Verfügbarkeit von Smartphones in der Bevölkerung eine besondere Rolle zugesprochen. Jedoch besteht seit längerem eine breite Diskussion hinsichtlich der noch eingeschränkten Evidenz von GuMA, welche in verschiedenen systematischen Übersichtsarbeiten als heterogen eingeschätzt wird (Agarwal, Lefevre & Labrique, 2017; Felix et al., 2018; Mathews et al., 2019). In einem systematischen Review von Bayambasuren et al. recherchierten die Autoren nach systematischen Reviews über randomisiert kontrollierte Studien (RCT) mit verschreibungsfähigen (Prescribable) GuMAs, die zum Download frei verfügbar waren (Byambasuren, Sanders, Beller & Glasziou, 2018). Verschreibungsfähige Apps definieren die Autoren als Apps, die derzeit verfügbar sind, die effektiv sind und die alleinstehend (stand-alone), ohne die Einbindung von Professionellen im Gesundheitswesen angewendet werden. Die Autoren schlossen insgesamt sechs systematische Reviews in ihre Analyse ein, aus denen sie wiederum 23 RCTs zu 22 unterschiedliche GuMAs extrahierten. In lediglich 11 dieser 23 RCTs konnte ein aussagekräftiger Effekt auf die Gesundheit oder auf ausgewählte Surrogatendpunkte nachgewiesen werden, die möglicherweise auf eine GuMA zurückzuführen waren. In der Gesamtbetrachtung war der Nachweis der Wirksamkeit jedoch von eingeschränkter Qualität, was gerade die Verschreibungsfähigkeit dieser Apps für Ärzte deutlich erschwert (Mathews et al., 2019). Ebenfalls monierten die Autoren sowohl bei den sechs eingeschlossenen systematischen Reviews als auch bei den 23 extrahierten RCTs ein hohes Risk of Bias, was ebenfalls zu einer Einschränkung der Ergebnisse beiträgt. Gleichermaßen finden sich in verschiedenen Anwendungsbereichen positive Ergebnisse hinsichtlich der Effektivität. Dies lässt sich etwa für das Selbstmanagement der Hypertonie (Li, Liang, Bu & Hesketh, 2020) oder das Selbstmanagement psychischer Krankheiten konstatieren (Linardon, Cuijpers, Carlbring, Messer & Fuller-Tyszkiewicz, 2019).

### Zusammenfassung

**Hintergrund:** Ähnlich der wenig qualitätsgesicherten Leitlinien Ende der 90er Jahre erlebt das deutsche Gesundheitssystem vor dem Hintergrund der Digitalisierung erneut eine Inflation an wenig qualitätsgesicherten Publikationen von digitalen Anwendungen. Vor diesem Hintergrund wird vielerorts eine schwache Evidenz von Gesundheits- und Medizin-Apps moniert. Um die Qualität von Studien und die Transparenz bei der Berichterstattung von Studien als Leistungserbringer und Anwendern einschätzen zu können, bedarf es gerade für Ärzte als auch Wissenschaftler einheitlicher Kriterien.

**Methodik:** Im September 2019 und ergänzend im Februar 2020 wurde eine Literaturrecherche in PubMed sowie Google Scholar durchgeführt. Gesucht wurden Frameworks und Reporting Guidelines, die bei Evaluationen zu Gesundheits- und Medizin-Apps angewendet werden können. Als Vergleichskriterien wurden Framework, Phasen/Items, Methodik und Fokus gegenübergestellt und explorativ ausgewertet.

**Ergebnis:** Es wurden sieben Publikationen zu Frameworks und Reporting Guidelines eingeschlossen. Diese unterscheiden sich in dem intendierten Fokus und dem Umfang sowie dem Inhalt der Bewertungskriterien. Die Frameworks und Reporting Guidelines adressieren Studien, die mit Gesundheits- und Medizin-Apps durchgeführt werden, wie z. B. das Mobile Health Evidence Reporting and Assessment (mERA) sowie den sog. DiGA-Leitfaden für das Fast-Track-Verfahren für digitale Gesundheitsanwendungen. Zudem sind methodenbezogene Reporting Guidelines hervorzuheben, welche primär einen methodischen Fokus auf die Anwendung legen, bspw. bei randomisiert kontrollierten Studien.

**Schlussfolgerung:** Angesichts der in dieser Arbeit dargestellten Optionen von Standards in der Berichterstattung von digitalen Anwendungen konnte der vom BfArM vorgestellte Leitfaden um weitere Items erweitert werden, um Leistungserbringern und Anwendern mehr Sicherheit zu geben und gleichzeitig Abstriche im Qualitätsanspruch zu verhindern.

### Schlüsselwörter

Digital Health, Gesundheits-Apps, eHealth, mHealth, Evaluation

### Crossref/doi

doi: <http://doi.org/10.24945/MVF.02.21.1866-0533.2300>

>> Die Gründe für die eingeschränkte Evidenz von GuMA sind vielfältig. Häufig werden die eingeschränkte methodische Qualität der Studien benannt (Tønning, Kessing, Bardram & Faurholt-Jepsen, 2019), hohe Drop-Out Raten, unspezifisch beschriebene Interventionen (Vilardaga, Casellas-Pujol, McClernon & Garrison, 2019) oder eine hohe Diversität der Interventionskomponenten einer GuMA bei einem gleichen Anwendungsfall, was die Evidenzsynthese deutlich erschwert (Thurnheer, Gravestock, Pichierr, Steurer, & Burgstaller, 2018). Des Weiteren ist hervorzuheben, dass die aus klinischer Perspektive empfohlenen Maßnahmen zur Therapie einer Erkrankung, die z. B. durch Leitlinien empfohlen werden, sich häufig nicht in

### Fünf Kernaussagen

- Mit der Digitalisierung im Gesundheitswesen sind große Erwartungen an die Qualitätssteigerung der gesundheitlichen Versorgung geknüpft, dies gilt gerade für den Einsatz von Gesundheits- und Medizin-Apps.
- Die Evidenz vieler Gesundheits- und Medizin-Apps ist derzeit noch eingeschränkt und die Bewertung der Qualität von Studien ist insbesondere für Ärzte und Forscher schwierig und intransparent.
- Frameworks und Reporting Guidelines stellen eine wertvolle Unterstützungsmöglichkeit dar, um Ärzte und Wissenschaftler bei der Bewertung der Qualität von Studien zu GuMA zu unterstützen.
- Frameworks und Reporting Guidelines finden kaum Anwendung bei Studien zu Gesundheits- und Medizin-Apps und vielerorts fehlt es an Wissen über deren Verfügbarkeit und Anwendung.
- Gerade vor dem Hintergrund des Digitale-Versorgung-Gesetz wird für Ärzte und Wissenschaftler der Bedarf zunehmend steigen, Studien zu Gesundheits- und Medizin-Apps in ihrer Qualität zu bewerten. Folglich ist es wichtig, gängige Frameworks und Reporting Guidelines anwenden zu können.

den Funktionalitäten von GuMA widerspiegeln. Demzufolge sind der Inhalt und die damit verbundenen Funktionalitäten von GuMAs an sich oftmals nicht evidenzbasiert. Hierfür finden sich zahlreiche Hinweise in systematischen Reviews zu verschiedenen Krankheitsbildern, wie beispielsweise in einem systematischen Review über GuMAs zur Unterstützung des Selbst-Managements von Diabetes (Adu, Malabu, Callander, Malau-Aduli & Malau-Aduli, 2018) oder in einem systematischen Review zu GuMAs zur Unterstützung des Schmerzmanagements (Thurnheer et al., 2018).

Dass sich die Diskussion hinsichtlich der Evidenz von GuMAs im deutschen Kontext deutlich intensiviert hat, ist besonders auf die Einführung des Digitale-Versorgung-Gesetz (DVG) zurückzuführen (EbM-Netzwerk, 2019; Krüger-Brand & E., 2019). Versicherte haben infolge des DVG nach § 33 a SGB V Anspruch auf die Versorgung mit Medizinprodukten niedriger Risikoklasse (I, IIa), deren medizinische Zweckbestimmung im wesentlichen mit digitalen Technologien vermittelt wird. Im Rahmen des DVG wird der Begriff der digitalen Gesundheitsanwendung (DiGA) verwendet. Das DVG fordert, dass erstattungsfähige digitale Gesundheitsanwendungen nach § 139e SGB V (2) den Anforderungen an Sicherheit, Funktionstauglichkeit und Qualität gemäß den Medizinprodukttrichtlinien MDD, MDR entsprechen sowie die Anforderungen an Datenschutz und Datensicherheit nach dem Stand der Technik erfüllen sowie positive Versorgungseffekte aufweisen. Dabei kann ein positiver Versorgungseffekt sowohl ein medizinischer Nutzen gemessen an patientenrelevanten Outcomeparametern als auch eine patientenrelevante Struktur- und Verfahrensverbesserung in der Versorgung sein.

Die Aufnahme von DiGA in das DiGA-Verzeichnis des Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) soll unter Berücksichtigung der Grundsätze der evidenzbasierten Medizin (EbM) erfolgen. In der EbM hat sich die Anwendung von Frameworks und Reporting Guidelines etabliert, die es ermöglichen, die Qualität und Transparenz des Reportings von Studien zu bewerten und die Replizierbarkeit der Ergebnisse zu verbessern. Für komplexe Interventionen stehen etwa das Framework des Medical Research Council (Craig et al., 2012) und die TIDieR Checkliste zur Verfügung (Hoffmann et al., 2014). Im Gegensatz zu Studien, in denen nicht-digitale Interventionen auf ihre Wirksamkeit geprüft werden, hat sich bei GuMAs die Anwendung von Frameworks und Reporting Guidelines noch nicht hinreichend etabliert (Agarwal et al., 2017; Vilardaga et al., 2019). Aufgrund der beschriebenen Problemlagen hinsichtlich des Reportings von Studien zu GuMAs und der eingeschränkten Evidenz soll der folgende Beitrag einen Überblick über derzeit verfügbare Frameworks und Reporting Guidelines bereitstellen.

## Methoden

Diese Arbeit basiert auf einer Literaturrecherche in PubMed und Google Scholar zwischen September 2019 und Februar 2020. Die Recherche basierte auf der Fragestellung: Welche Frameworks und Reporting Guidelines existieren für das Reporting von Studienergebnissen mit GuMA? Es wurden Publikationen zu Frameworks und Re-

porting Guidelines sowie zu Checklisten von GuMA eingeschlossen. Eingeschlossen wurden Publikationen in deutscher und englischer Sprache seit 2007 (Markteinführung des ersten iPhones). Für die Recherche wurden die folgenden Suchbegriffe verwendet: „(Digital Health OR eHealth OR mHealth OR Digital Health Intervention OR Telemedicine OR Teleradiology OR Telemonitoring OR Telehealth) AND (Evaluation OR Framework OR Checklist OR Standards OR Methods)“. Im Rahmen der Recherche erfolgte eine kontinuierliche Rückkopplung der Autoren hinsichtlich des Ein- oder Ausschlusses möglicher Publikationen.

## Ergebnisse

Es wurden sieben Publikationen zu Frameworks und Reporting Guidelines identifiziert. Die Ergebnisdarstellung adressiert im Folgenden jeweils die Darstellung von (1) Name und Ziel des Frameworks, (2) Autoren oder Entwickler, (3) Methodisches Vorgehen bei der Entwicklung, (4) Fokus, (5) Phasen, Kategorien oder Items, (6) Weitere Spezifika. Eine vergleichende Darstellung der Ergebnisse findet sich in Tabelle 4.

### (1) Mobile Health (mHealth) Evidence Reporting and Assessment Checklist

Die Mobile Health Evidence Reporting and Assessment (mERA) Checkliste wurde zur Steigerung der methodischen Qualität und der Transparenz bei dem Reporting von mHealth-Interventionen durch die Technical Evidence Review Group (mTERG) der World Health Organization (WHO) entwickelt (Agarwal et al., 2017). Die Entwicklung der mERA erfolgte anhand der Methodik zur Entwicklung von Reporting Guidelines angelehnt an Moher et al. (Moher, Schulz, Simera & Altman, 2010) welche sich in drei Schritte unterteilt:

- (1) Einberufung einer Expertenarbeitsgruppe,
- (2) Einberufung eines globalen Expertenprüfungsgremiums
- (3) Pilottestung der Checkliste (Agarwal et al., 2016).

Die mREA beinhaltet zwei zentrale Komponenten: Die (I) mREA Checkliste, die sich aus 16 Items zusammensetzt (Inhalt, Kontext und technische Umsetzung einer mHealth Intervention) (siehe Abbildung 1) und eine (II) Methoden Checkliste, die sich aus 29 Items zusammensetzt (Gordon et al., 2020). Die mERA ist unabhängig von dem Studiendesign einzusetzen, in dessen Rahmen die Evaluation einer GuMA stattfindet. Deswegen empfehlen die Autoren, mREA in Kombination mit Checklisten für das jeweilige Studiendesign zu ver-

Unterpunkte zur Erweiterung von CONSORT-Item 5 (Beschreibung der Intervention)	
• Infrastruktur	• Kostenbewertung
• Technologische Plattform	• Adoptionseingaben/Programmeingabe
• Interoperabilität im Kontext von Gesundheitssystemen (HIS)	• Einschränkungen der Interventionstreue
• Interventionstreue („delivery“)	• Kontextuelle Anpassungsfähigkeit
• Inhalt der Intervention	• Reproduzierbarkeit
• Benutzerfreundlichkeit/Inhaltstests	• Datensicherheit
• Benutzer-Feedback	• Einhaltung nationaler Richtlinien oder gesetzlicher Bestimmungen
• Zugang zu den Teilnehmern	• Grad der Abweichung vom Studienprotokoll / Treue der Intervention („fidelity“)

Tab. 1: Mobile Health (mHealth) Evidence Reporting and Assessment (16).

Schritt 1 – Vorhergehende Überlegungen	
• Zweck der telemedizinischen Anwendung	• Internationale, regionale oder lokale Bewertung
• Relevante Alternativen	• Entwicklungsgrad der Anwendung
Schritt 2 – Multidisziplinäre Bewertung	Schritt 3 – Beurteilung der Übertragbarkeit
• Gesundheitsproblem und Eigenschaften der telemedizinischen Anwendung	• Grenzübergreifend
• Sicherheit	• Skalierbarkeit
• Klinische Effektivität	• Generalisierbarkeit
• Patientenperspektive	
• Ökonomische Aspekte	
• Organisatorische Aspekte	
• Soziokulturelle, ethische und juristische Aspekte	

Tab. 2: Modell zur Evaluation von telemedizinischen Applikationen (23).

wenden, wie etwa CONSORT für randomisiert kontrollierte Studien (RCT) oder STROBE für Beobachtungsstudien (Agarwal et al., 2016).

**(2) The Model for Assessment of Telemedicine**

Das Model for Assessment of Telemedicine (MAST) wurde als Framework entwickelt, um die Effektivität und den Beitrag zur Qualität der Versorgung einer telemedizinischen Anwendung einzuschätzen (Kidholm et al., 2012). MAST wurde zwar nicht primär für den Einsatz von Studien zu GuMA entwickelt, jedoch können GuMA gerade in der Telemedizin eine Komponente einer Intervention sein. Ebenso wurde MAST vordergründig nicht für das Reporting von Studienergebnissen entwickelt (Kidholm et al., 2012). MAST kann jedoch überaus hilfreich sein, um die Qualität des Reportings von telemedizinischen Studien einzuschätzen (Abbey et al., 2018). Das Modell wurde durch Workshops mit Nutzern und Stakeholdern im Bereich Telemedizin und durch ein systematisches Review (Ekelend, Bowes & Signe, 2010) entwickelt (Kidholm et al., 2012). Im Jahre 2016 wurde das Modell mit einer weiteren Delphi Umfrage validiert (Kristian, K., Tue, Emilie, & B., 2018). MAST umfasst drei Schritte (I) Vorhergehende Überlegungen (II) Multidisziplinäre Bewertung, (III) Bewertung der Übertragbarkeit der Ergebnisse (Kidholm, Clemensen, Caffery & Smith, 2017; Kidholm et al., 2012) (siehe Tab. 2). MAST ist in verschiedenen Sprachen verfügbar, auch in einer deutschsprachigen Version (Allner, Wilfling, Kidholm, & Steinhäuser, 2019). In der gängigen Forschungspraxis zeigt sich, dass aus MAST häufig nur einzelne der drei Schritte angewendet werden. Die Empfehlung ist jedoch, alle Schritte von MAST bei der Entwicklung und Evaluation von telemedizinischen Interventionen anzuwenden (Kidholm et al., 2017).

**(3) Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and online TeleHealth**

Das Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and online TeleHealth (CONSORT-EHEALTH) wurde entwickelt, um die Qualität des Reporting von RCTs im Bereich eHealth oder mHealth zu verbessern (Eysenbach &

Group, 2011). Die Items des CONSORT-EHEALTH wurden, orientiert an der originalen Consolidated Standards of Reporting Trials (CONSORT) Checkliste (Hoffmann et al., 2014) für das Reporting von RCTs, entwickelt. Hierzu erfolgte eine systematische Literaturrecherche sowie eine anschließende Expertenumfrage und einen Expertenworkshop.

CONSORT-EHEALTH besteht aus 27 Items, die wiederum teilweise in Sub-Items aufgliedert sind. Summiert man die Items und Sub-Items, umfasst

CONSORT-EHEALTH insgesamt 52 Items bzw. Sub-Items. Von diesen 52 Items bzw. Sub-Items werden für das Reporting von RCTs 17 als „essential“ und 35 als „highly recommended“ ausgewiesen. Im Besonderen wurde das Item 5 „Beschreibung der Intervention“ der originalen CONSORT-Checkliste spezifisch an die Eigenschaften von eHealth und mHealth Interventionen adaptiert (siehe Tab. 3).

**(4) Integrate, Design, Assess and Share Framework for developing digital Health Behavior Change Interventions – IDEAS**

Das Integrate, Design, Assess, and Share Framework (IDEAS) wurde spezifisch für die Entwicklung und Evaluation von digitalen Interventionen zur Veränderung des Gesundheitsverhaltens entwickelt (Ann, N, C, D & Sutton, Stephen, 2016). IDEAS kombiniert die Stärken einer verhaltenstheoretischen Betrachtung, des Design Thinkings, einer Evaluation sowie der Dissemination und bildet ein Rahmenwerk für die Entwicklung von digitalen Intervention zur Verhaltensänderung. Das IDEAS-Framework besteht aus 10 Phasen: (1) Hineinversetzen in die Zielgruppe, (2) Spezifizierung des adressierten Verhaltens, (3) Grundlagen der Verhaltenstheorie, (4) Ausarbeitung von Implementierungsstrategien, (5) Entwicklung von Prototypen eines möglichen (digitalen) Produktes, (6) Einholen von

Erweiterung zur Beschreibung der E-Health-Intervention
• Namen, Referenzen, Zugehörigkeit der Entwickler, Sponsoren und Eigentümer (Sehr empfehlenswert)
• Geschichte und Entwicklung (Sehr empfehlenswert)
• Überarbeitungen und Aktualisierungen (Sehr empfehlenswert)
• Informationen über Qualitätssicherungsmethoden (Sehr empfehlenswert)
• Sicherstellen der Reproduzierbarkeit durch Veröffentlichung des Quellcodes (vorzugsweise als Open Source) und/oder Bereitstellung von Screenshots/Bildschirmabbildungsvideos und/oder Bereitstellung von Flussdiagrammen der verwendeten Algorithmen (Sehr empfehlenswert)
• Nachhaltige digitale Verfügbarkeit, wenn zutreffend (z. B. durch eine URL) (Sehr empfehlenswert)
• Zugang zur Intervention (Essenziell)
• Art und Weise der Durchführung, Eigenschaften/Funktionalitäten/Komponenten der Intervention und des Komparators sowie Beschreibung des theoretischen Rahmens (Essenziell)
• Nutzungsparameter beschreiben (Sehr empfehlenswert), Grad der menschlichen Beteiligung beschreiben (Sehr empfehlenswert)
• Alle verwendeten Eingabeaufforderungen/Erinnerungen melden (Essenziell)
• Etwaige zusätzliche Interventionen (einschließlich Schulung/Unterstützung) beschreiben (Essenziell)

Tab. 3: Adaptierung bzw. Erweiterung des Items 5 der originalen CONSORT-Checkliste.



Feedback der Benutzer, (7) Erstellung eines „minimal viable product“, (8) Pilotierung Wirksamkeit und Usability, (9) Bewertung der Wirksamkeit durch eine RCT und (10) Austausch zur Intervention und den Ergebnissen. Die im IDEAS Framework beinhaltenden Phasen werden in 4 übergeordnete Kategorien eingeordnet: Integrieren, Entwerfen, Bewerten und Teilen (siehe Abb. 4).

### (5) Design and Evaluation Framework for Digital Health Interventions – DEDHI

Das Design and Evaluation Framework for Digital Health Interventions (DEDHI) stellt einen Überblick über zentrale Evaluationskriterien und Barrieren bei der Implementation von digitalen Interventionen zu gesundheitsbezogenen Themen bereit. DEDHI bezieht sich sowohl auf die Evaluation von Smartphone-Interventionen als auch auf Interventionen im Bereich Digital Health allgemein. Durch die Autoren wird allgemein der Begriff „Digital Health Interventions (DHIs)“ verwendet. Zur Entwicklung von DEDHI wurden eine systematische Literaturrecherche sowie eine qualitative Inhaltsanalyse der Evaluationskriterien durchgeführt (Kowatsch, Otto, Harperink, Cotti, & Schlieter, 2020). Im Anschluss daran wurden vier übergeordneten Phasen gebildet, die an die Multiphase Optimization Strategy (MOST) angelehnt wurden (Collins, 2018; Kowatsch et al., 2020) DEDHI besteht aus den vier Phasen (1) Vorbereitung, (2) Optimierung, (3) Evaluation, und (4) Implementation. Alle vier Phasen beinhalten jeweils vier Unterkategorien (I) Ziele und Aufgaben, (II) Technische Reife, (III) Evaluationskriterien (IV) Barrieren bei der Implementierung (Kowatsch et al., 2020). Eine Validierung von DEDHI fand nach Angaben der Autoren noch nicht statt (Kowatsch et al., 2020).

### (6) The Evidence in Electronic Health (eHealth) Evaluation Model

Das Evidence in Electronic Health (eHealth) Evaluation Model Design (EEHEM) wurde von Enam et al. (Enam, Torres-Bonilla & Eriksson, 2018) im Rahmen eines systematischen Reviews entwickelt. Das Review der Autoren hatte in erster Linie das Ziel, Phasen zu identifizieren, in denen die Evaluation von eHealth-Interventionen durchgeführt wird und zudem Aspekte hinsichtlich der Effektivität und Effizienz zu beschreiben. Die Autoren umfassen mit eHealth-Interventionen keine spezifische Art der Technologie, wie etwa GuMA. Bei dem EEHEM handelt es sich folglich um ein eher übergeordnetes Model, welches für verschiedene Formen digitaler Interventionen eingesetzt werden kann. Aus den Ergebnissen dieses systematischen Reviews entwickelten die Autoren das EEHEM, welches aus den sechs Phasen besteht:

- (1) Design,
- (2) Pre-Testung,
- (3) Pilotstudien,
- (4) Pragmatische Studie,
- (5) Evaluation und
- (6) Postintervention.

Eingeschlossene Frameworks				
	Framework	Phasen/Items		Fokus
1.	mERA	16 Items	Entwicklung von Reporting Guidelines nach Moher, Schulz, Simera und Altman (2010)	mHealth
2.	MAST	3 Schritte	Workshops	Telemedicine
3.	CONSORT-EHEALTH	53 Items	Delphi Umfrage	eHealth RCTs
4.	IDEAS	10 Phasen	Workshops	Interventionen mit Verhaltenstheoretischem Hintergrund
5.	DEDHI	Vier Phasen/vier Unterkategorien	n. a.	Digital Health Interventions (DHIs)“
6.	EEHEM	6 Phasen	Systematische Literaturrecherche/Inhaltsanalyse	eHealth Interventionen
7.	NICE	3 Evidence Tiers	n.a.	Digital Health Technologies (DHT)

Tab. 4: Übersicht der eingeschlossenen Frameworks.

Die Autoren heben hervor, dass eine robuste Evaluation sowie der Nachweis von Effektivität und Effizienz einer eHealth-Intervention nur möglich sind, wenn alle Phasen des EEHEM bei der Evaluation berücksichtigt werden. In der Publikation der Autoren finden sich keine Angaben zur Validierung des Modells.

### (7) Evidence Standards Framework for Digital Health Technologies

Das Evidence Standards Framework for Digital Health Technologies des National Health System (NHS) England beschreibt Standards zur Entwicklung der Evidenz digitaler Technologien sowie Standards zur Entwicklung der Evidenz ökonomischer Evaluationen für digitale Technologien (NICE, 2019). Digitale Technologien werden in dem NICE-Framework anhand ihrer funktionalen Eigenschaften und dem mit der Anwendung einhergehenden Risiko für die Nutzer verschiedenen Evidenzstufen zugeordnet: Je höher die Evidenzstufe, desto höhere Anforderungen werden an das notwendige Studiendesign für den Nutznachweis gestellt. In diesem Zusammenhang wird argumentiert, dass digitale Technologien wie andere Interventionen um die begrenzten Ressourcen im Gesundheitswesen konkurrieren, infolgedessen sollten die gleichen Anforderungen an die Evidenz gestellt werden, wie dies bei nicht-digitalen Interventionen der Fall sei (Sauerland, 2019).

Die Evidenzstufen wurden anhand der WHO-Klassifizierung digitaler Technologien gebildet (WHO, 2020):

**Evidenzstufe I:** Digitale Technologien, die für organisatorische Aspekte des Gesundheitswesens Vorteile bewirken, jedoch keinen direkten gesundheitlichen Nutzen für die Nutzer bewirken, wie z. B. elektronische Verschreibungssysteme oder elektronische Patientenakten. Auf dieser Evidenzstufe wird gefordert, dass Praktikabilität und Nutzerakzeptanz evaluiert werden (NICE, 2019; Sauerland, 2019)

**Evidenzstufe II:** Diese Stufe beinhaltet drei Funktionalitäten: Informieren, einfaches Monitoring von Gesundheitsdaten und Kommunikation. Digitale Technologien im Bereich Informieren stellen einfache Gesundheitsinformationen bereit, wie z. B. Informationen über eine Krankheit und die damit verbundenen Behandlungen. Der Bereich Einfaches Monitoring beinhaltet digitale Technologien, mit

denen einfache Gesundheitsparameter archiviert werden können, um etwa Gesundheitstagebücher zu erstellen. Ausgeschlossen sind digitale Technologien, durch die Gesundheitsparameter mit weiteren Personen geteilt werden können, wie etwa mit Ärzten. Im Bereich Kommunizieren sind digitale Technologien inbegriffen, die eine Zwei-Wege-Kommunikation zwischen Anwendern und Experten ermöglicht. Dies impliziert z. B. Instant-Messaging-Apps oder Konsultationssoftware für Videokonferenzen. Auf dieser Evidenzstufe wird gefordert, dass Praktikabilität und Nutzerakzeptanz evaluiert und eine Qualitätssicherung der Gesundheitsinformationen erfolgt (NICE, 2019; Sauerland, 2019)

**Evidenzstufe IIIa:** Digitale Technologien, die zur Prävention und Behandlung von Krankheiten eingesetzt werden oder die bei Personen mit einer diagnostizierten Erkrankung auf das Selbstmanagement der Erkrankung abzielen. Auf der Evidenzstufe IIIa wird eine Beobachtungsstudie oder eine quasi Experimentelle Studie zur Evaluation gefordert (NICE, 2019).

**Evidenzstufe IIIb:** Digitaler Technologien, mit denen aktiv die Behandlung unterstützt oder das aktive Monitoring einer Erkrankung durchgeführt wird. In dieser Stufe muss der Nutzen mit einer randomisiert kontrollierten Studie nachgewiesen werden. Auf der Evidenzstufe IIIb wird ein Studiendesign in Form eines RCT gefordert. Drei Aspekte für die Einordnung des NICE-Framework sind zu berücksichtigen. Zum einen erfolgt die Klassifizierung einer digitalen Intervention unabhängig davon, ob ein CE-Kennzeichnung vorliegt oder ob es sich bei einer Intervention um ein Medizinprodukt handelt (NICE, 2019). Jedoch geht das Framework davon aus, dass digitale Interventionen auf Stufe 3b mit den Regularien für Medizinprodukte korrelieren. Zum anderen werden digitale Interventionen ausgeschlossen, die adaptive Algorithmen beinhalten, welche sich folglich während der Anwendung verändern können (NICE, 2019).

### Evaluation im Kontext des Digitale-Versorgung-Gesetz

Abschließend soll an dieser Stelle noch ein Bezug zur Situation in Deutschland hinsichtlich des Digitale-Versorgung-Gesetz (DVG) genommen werden. Im deutschen Kontext hat das Thema der Evaluation von digitalen Interventionen im Allgemeinen und GuMA im Speziellen eine besondere Dynamik und eine neue Tiefe der wissenschaftlichen Diskussion entwickelt. Dies ist insbesondere auf die Einführung des DVG am 19. Dezember 2019 zurückzuführen (BfArM, 2020a). Abgrenzend zu der in diesem Artikel durchgeführten Recherche soll aufgrund der Besonderheit für den deutschen Anwendungskontext beschrieben werden, welche Aspekte im Rahmen des DVG an die Ausgestaltung und Evaluation einer DiGA geknüpft sind. Innerhalb dieser Gesetzgebung haben Versicherte den gesetzlichen Anspruch auf die „App auf Rezept“, welche nach dem § 33 a SGB V z. B. durch Ärzte und Psychotherapeuten verschrieben werden können. Als eine Besonderheit definiert das DVG die „App auf Rezept“ als sog. Digitale Gesundheitsanwendung (DiGA). Nach Definition des DVG sind DiGA Smartphone-Apps, die als Medizinprodukte mit einer niedrigen Risikoklasse (I & IIa) ausgewiesen sind und deren Hauptfunktion im wesentlichen auf digitalen Technologien basieren (BfArM, 2020b). Nach § 139e SGB V (2) müssen DiGA verschiedene Anforderungen an die Sicherheit, Funktionstauglichkeit und Qualität von Medizinprodukten sowie an Datenschutz und Datensicherheit nach dem aktuellen Stand der Technik erfüllen. Des Weiteren ist ein wesentlicher Aspekt, dass DiGA positive Versorgungseffekte

aufweisen müssen. Solche positiven Versorgungseffekte können sowohl ein medizinischer Nutzen sein als auch patientenrelevante Struktur- und Verfahrensverbesserungen, die in der Versorgung eine Rolle spielen können. Es wird im DVG explizit auf die Berücksichtigung von Grundsätzen der evidenzbasierten Medizin (EbM) hingewiesen. Eine wesentliche Voraussetzung dafür, dass eine GuMA als DiGA als „App auf Rezept“ durch Ärzte und Psychotherapeuten verschrieben werden kann, ist, dass ein als Fast-Track bezeichnetes Prüfverfahren erfolgreich abgeschlossen werden muss, welches durch das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) durchgeführt wird. Nach erfolgreichem Abschluss des Prüfverfahrens wird die entsprechende DiGA in ein beim BfArM verwaltetes Verzeichnis für erstattungsfähige digitale Gesundheitsanwendungen aufgenommen (§ 139e SGB V). Um in dieses Verzeichnis aufgenommen zu werden, müssen durch die Hersteller verschiedene Anforderungen erfüllt sein. Zudem ist es eine wesentliche Grundlage, dass die Anforderungen an das Medizinproduktegesetz erfüllt werden müssen (BfArM, 2020b). Eine Übersicht der wesentlichen Anforderungen, die DiGA erfüllen müssen, findet sich in Tabelle 5.

### Diskussion

Im Rahmen dieser Recherche konnten insgesamt sieben unterschiedliche Frameworks und Reporting Guidelines identifiziert werden, die Forscher und Kliniker dabei unterstützen können, die Qualität des Reportings von Studien bei GuMA einzuschätzen. Des Weiteren wurde über ausgewählte Aspekte des DVG berichtet, die im deutschen Anwendungskontext eine Rolle spielen können, je-

Anforderungen an DiGA
<b>Robustheit</b>
<ul style="list-style-type: none"> <li>• Verwendung der DiGA möglichst ohne Störungen, Datenverluste, Übertragungsfehler oder Schwierigkeiten bei der Verbindung mit Geräten, sowie fehlerhafte Dateneingaben oder -übernahmen</li> </ul>
<b>Verbraucherschutz</b>
<ul style="list-style-type: none"> <li>• Transparenz gegenüber den Nutzern in Bezug auf die Zweckbestimmung und die Funktionalität der DiGA.</li> </ul>
<b>Nutzerfreundlichkeit</b>
<ul style="list-style-type: none"> <li>• DiGA muss intuitiv nutzbar und erlernbar sein. Ab dem 1. Januar 2021 müssen alle gelisteten DiGA entweder Bedienhilfen für Menschen mit Einschränkungen beinhalten oder die durch die Plattform angebotenen Bedienhilfen unterstützen.</li> </ul>
<b>Unterstützung der Leistungserbringer</b>
<ul style="list-style-type: none"> <li>• Anfragen durch die Nutzer müssen zeitnah beantwortet werden.</li> </ul>
<b>Qualität der medizinischen Inhalte</b>
<ul style="list-style-type: none"> <li>• Inhalte der DiGA müssen auf einer medizinisch-fachlichen Grundlage basieren sowie Bereitstellung geeigneter Gesundheitsinformationen für die jeweilige Zielgruppe.</li> </ul>
<b>Patientensicherheit</b>
<ul style="list-style-type: none"> <li>• Risiken müssen bei der Anwendung möglichst gering gehalten werden.</li> </ul>
<b>Nachweis positiver Versorgungseffekte</b>
<ul style="list-style-type: none"> <li>• Verbesserung des Gesundheitszustands</li> <li>• Verkürzung der Krankheitsdauer</li> <li>• Verlängerung des Überlebens</li> <li>• Verbesserung der Lebensqualität</li> <li>• Patientenrelevante Struktur- und Verfahrensverbesserungen (pSVV)</li> </ul>

Tab. 5: Anforderungen an DiGA.

doch im Vergleich zu den anderen hier vorgestellten Inhalten kein Framework oder eine Reporting Guideline darstellt.

Die Frameworks und Reporting Guidelines unterscheiden sich in dem jeweiligen Fokus und in der Anzahl der Items, mit denen eine Einschätzung vorgenommen werden soll. Vergleicht man die unterschiedlichen Frameworks und Reporting Guidelines, so ist

festzustellen, dass die Begrifflichkeiten für digitale Interventionen durchaus unterschiedlich verwendet werden. Beispiele sind hierbei u. a. mHealth-Interventionen, Digital Health Technologies, eHealth Interventionen oder Gesundheits- und Medizin-Apps. Es ist davon auszugehen, dass es aufgrund der Pluralität dieser Begrifflichkeiten bei Ärzten und Forschern sowie bei weiteren Stakeholdern bei der

## Literatur

- Abbey, E., Maiken, H., Camilla, V., Lyngsie, H., Lisa, Utoft, N., Carsten, Kirsten, B., & Helle, P. (2018). Cocreated Smartphone App to Improve the Quality of Life of Adolescents and Young Adults with Cancer (Kræftværket): Protocol for a Quantitative and Qualitative Evaluation. *JMIR Res Protoc*, 7(5), e10098. doi: 10.2196/10098
- Adu, M. D., Malabu, U. H., Callander, E. J., Malau-Aduli, A. E., & Malau-Aduli, B. S. (2018). Considerations for the Development of Mobile Phone Apps to Support Diabetes Self-Management: Systematic Review. *JMIR MHealth and UHealth*, 6(6), e10115. doi: 10.2196/10115
- Agarwal, S., Lefevre, A. E., & Labrique, A. B. (2017). A Call to Digital Health Practitioners: New Guidelines Can Help Improve the Quality of Digital Health Evidence. *JMIR MHealth and UHealth*, 5(10), e136. doi: 10.2196/mhealth.6640
- Agarwal, S., LeFevre, A. E., Lee, J., L'Engle, K., Mehl, G., Sinha, C., ... Group, W. mHealth T. E. R. (2016). Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. *BMJ*, 352, i1174. doi: 10.1136/bmj.i1174
- Ahmadiankalati, M., Steins-Loeber, S., & Paslakis, G. (2020). Review of Randomized Controlled Trials Using e-Health Interventions for Patients With Eating Disorders. *Frontiers in Psychiatry*, 11, 568. doi: 10.3389/fpsyt.2020.00568
- Allner, R., Wilfling, D., Kidholm, K., & Steinhäuser, J. (2019). Telemedizinprojekte im ländlichen Raum Deutschlands. Eine systematische Bewertung mit dem „Modell zur Evaluation von telemedizinischen Anwendungen“. *Zeitschrift Für Evidenz, Fortbildung Und Qualität Im Gesundheitswesen*, 141-142, 89-95. doi: 10.1016/j.zefq.2019.03.005
- Ann, M., Sarah, N. R., Thomas, C. K., Abby, D. G., Christopher, & Sutton, Stephen. (2016). IDEAS (Integrate, Design, Assess, and Share): A Framework and Toolkit of Strategies for the Development of More Effective Digital Interventions to Change Health Behavior. *J Med Internet Res*, 18(12), e317. doi: 10.2196/jmir.5927
- BfArM (2020a) Bundesinstitut für Arzneimittel und Medizinprodukte. Digitale Gesundheitsanwendungen. In [https://www.bfarm.de/DE/Medizinprodukte/DVG/\\_node.html](https://www.bfarm.de/DE/Medizinprodukte/DVG/_node.html) (Abgerufen am 20.05.2020)
- BfArM. (2020b). Das Fast-Track-Verfahren für digitale Gesundheitsanwendungen (DiGA) nach § 139e SGB V. Ein Leitfaden für Hersteller, Leistungserbringer und Anwender. In [https://www.bfarm.de/DE/Medizinprodukte/DVG/\\_node.html](https://www.bfarm.de/DE/Medizinprodukte/DVG/_node.html) (Abgerufen am 20.05.2020)
- Byambasuren, O., Sanders, S., Beller, E., & Glasziou, P. (2018). Prescribable mHealth apps identified from an overview of systematic reviews. *Npj Digital Medicine*, 1(1), 12. doi: 10.1038/s41746-018-0021-9
- Collins, L. (2018). Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimization Strategy (MOST). Springer International Publishing.
- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2012). Developing and evaluating complex interventions: the new Medical Research Council guidance. *International Journal of Nursing Studies*, 50(5), 587-592. doi: 10.1016/j.ijnurstu.2012.09.010
- EbM-Netzwerk. (2019). Stellungnahme des Deutschen Netzwerks Evidenzbasierte Medizin zum Entwurf eines Gesetzes für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale Versorgung-Gesetz - DVG) vom 15.5.2019.
- Ekeland, A., Bows, A., & Signe, F. (2010). Effectiveness of telemedicine: A systematic review of reviews. *International Journal of Medical Informatics*, 79, 736-771. doi: <https://doi.org/10.1016/j.ijmedinf.2010.08.006>
- Enam, A., Torres-Bonilla, J., & Eriksson, H. (2018). Evidence-Based Evaluation of eHealth Interventions: Systematic Literature Review. *Journal of Medical Internet Research*, 20(11), e10971. doi: 10.2196/10971
- Eysenbach, G., & Group, C.-E. (2011). CONSORT-EHEALTH: Improving and Standardizing Evaluation Reports of Web-based and Mobile Health Interventions. *Journal of Medical Internet Research*, 13(4), e126. doi: 10.2196/jmir.1923
- Felix, G., Indra, J., Mark, C., Samantha, R., Neelam, P., & John, P. (2018). What is an appropriate level of evidence for a digital health intervention? *The Lancet*, 392(10165), 2665-2667. doi: 10.1016/s0140-6736(18)33129-5
- Gordon, W. J., Landman, A., Zhang, H., & Bates, D. W. (2020). Beyond validation: getting health apps into clinical practice. *Npj Digital Medicine*, 3(1), 14. doi: 10.1038/s41746-019-0212-z
- Hennessy, M., Heary, C., Laws, R., Rhoon, L., Toomey, E., Wolstenholme, H., & Byrne, M. (2019). The effectiveness of health professional-delivered interventions during the first 1000 days to prevent overweight/obesity in children: A systematic review. *Obesity Reviews*, 20(12), 1691-1707. doi: 10.1111/obr.12924
- Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., ... Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*, 348(mar07 3), g1687-g1687. doi: 10.1136/bmj.g1687
- Kidholm, K., Clemensen, J., Caffery, L. J., & Smith, A. C. (2017). The Model for Assessment of Telemedicine (MAST): A scoping review of empirical studies. *Journal of Telemedicine and Telecare*, 23(9), 803-813. doi: 10.1177/1357633x17721815
- Kidholm, K., Ekeland, A. G., Jensen, L. K., Rasmussen, J., Pedersen, C. D., Bows, A., ... Bech, M. (2012). A MODEL FOR ASSESSMENT OF TELEMEDICINE APPLICATIONS: MAST. *International Journal of Technology Assessment in Health Care*, 28(1), 44-51. doi: 10.1017/s0266462311000638
- Kowatsch, T., Otto, L., Harperink, S., Cotti, A., & Schlieter, H. (2020). A design and evaluation framework for digital health interventions. *IT - Information Technology*, 61(5-6), 253-263. doi: 10.1515/itit-2019-0019
- Kristian, K., K., J., Lise, Tue, K., Emilie, N., & B., H., Mette. (2018). Validity of the Model for Assessment of Telemedicine: A Delphi study. *J Telemed Telecare*, 24(2), 118-125. doi: 10.1177/1357633x16686553
- Krüger-Brand, & E., H. (2019). Digitale Versorgung-Gesetz: Im Detail noch nachschärfen. *Dtsch Arztebl.*
- Li, R., Liang, N., Bu, F., & Hesketh, T. (2020). The Effectiveness of Self-Management of Hypertension in Adults Using Mobile Health: Systematic Review and Meta-Analysis. *JMIR MHealth and UHealth*, 8(3), e17776. doi: 10.2196/17776
- Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., & Fuller Tyszkiewicz, M. (2019). The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. *World Psychiatry*, 18(3), 325-336. doi: 10.1002/wps.20673
- Mathews, S. C., McShea, M. J., Hanley, C. L., Ravitz, A., Labrique, A. B., & Cohen, A. B. (2019). Digital health: a path to validation. *NPJ Digital Medicine*, 2(1), 38. doi: 10.1038/s41746-019-0111-3
- Moher, D., Schulz, K. F., Simera, I., & Altman, D. G. (2010). Guidance for developers of health research reporting guidelines. *PLoS Medicine*, 7(2), e1000217. doi: 10.1371/journal.pmed.1000217
- NICE. (2019). Evidence standards framework for digital health technologies. In: <https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies> (Abgerufen am 10.12.2020)
- Sauerland, S. (2019). Digitale Gesundheitsinterventionen - Wann braucht es welche Evidenz? *KVH-Journal*, 4/2019.
- Thurnheer, S. E., Gravestock, I., Pichiari, G., Steurer, J., & Burgstaller, J. M. (2018). Benefits of Mobile Apps in Pain Management: Systematic Review. *JMIR MHealth and UHealth*, 6(10), e11231. doi: 10.2196/11231
- Tønning, M. L., Kessing, L. V., Bardram, J. E., & Faurholt-Jepsen, M. (2019). Methodological Challenges in Randomized Controlled Trials on Smartphone-Based Treatment in Psychiatry: Systematic Review. *Journal of Medical Internet Research*, 21(10), e15362. doi: 10.2196/15362
- Vilardaga, R., Casellas-Pujol, E., McClernon, J. F., & Garrison, K. A. (2019). Mobile Applications for the Treatment of Tobacco Use and Dependence. *Current Addiction Reports*, 6(2), 86-97. doi: 10.1007/s40429-019-00248-0
- Wang, Y., Min, J., Khuri, J., Xue, H., Xie, B., Kaminsky, L. A., & Cheskin, L. J. (2020). Effectiveness of Mobile Health Interventions on Diabetes and Obesity Treatment and Management: Systematic Review of Systematic Reviews. *JMIR MHealth and UHealth*, 8(4), e15400. doi: 10.2196/15400
- WHO. (2020). mHealth: new horizons for health through mobile technologies: second global survey on eHealth. *Global observatory for eHealth Series*, 3.

Studieninterpretation zu Verwirrungen kommen kann, wenn Studien eingeschätzt werden sollen. Gerade vor dem Hintergrund des DVG ist zu erwarten, dass die Anforderungen an die Interpretation von Studien zu digitalen Interventionen, insbesondere GuMA steigen wird. Infolgedessen ist es von hoher Bedeutung, auch Frameworks und Reporting Guidelines, die für den Bereich digitaler Interventionen entwickelt wurden, anzuwenden. Hierfür wäre es jedoch äußerst wünschenswert zu analysieren, wie konkret die hier identifizierten Frameworks und Reporting Guidelines in der wissenschaftlichen Praxis angewendet werden und welche Erfahrungen bei den Anwendern in Bezug auf den Nutzen bestehen. Es wäre zudem wünschenswert, dass hierzu weitere Forschung vorgenommen wird, um die Auswirkungen der Anwendung der Frameworks und Reporting Guidelines auf die Qualität der Studien zu untersuchen. <<

### Zitationshinweis

Kernebeck et al.: „Bewerten, aber wie? – Kriterien für die Evaluation und das Reporting von Studien zu Gesundheits- und Medizin-Apps“, in: „Monitor Versorgungsforschung“ (02/21), S. 69-75. doi: <http://doi.org/10.24945/MVF.02.21.1866-0533.2300>

### Autorenerklärung

Sven Kernebeck ist Mitarbeiter der Universität Witten/Herdecke in Witten. Ursula Kramer, Marcus Redaelli und Horst Christian Vollmar haben unentgeltlich an der Auswertung der Daten sowie der Interpretation der Ergebnisse mitgearbeitet. Die Analyse erfolgte ohne finanzielle Unterstützung.

## Evaluate and Report, but how? – Criteria for evaluation and reporting of studies on health and medical apps

**Background:** The quality of reporting of study results in the field of digital health is currently heterogeneous and in several cases poor and incomplete. Against this background, weak evidence of health and medical apps is criticized in many places. In order to be able to assess the quality of studies and the transparency in reporting studies as health care providers and users, uniform criteria are needed, especially for physicians as well as scientists.

**Methods:** A literature search was conducted in PubMed as well as Google Scholar in September 2019 and supplementally in February 2020. Frameworks and Reporting Guidelines that can be applied to evaluations of health and medical apps were included. Framework, phases/items, methodology, and focus were compared and exploratively evaluated as comparison criteria.

**Results:** Seven publications on frameworks and reporting guidelines were included. These differ in the intended focus and the scope and content of the evaluation criteria. The frameworks and reporting guidelines address studies conducted with health and medical apps, such as the Mobile Health Evidence Reporting and Assessment (mERA) and the so-called DiGA guide for the fast-track process for digital health apps. In addition, method-related reporting guidelines should be highlighted, which primarily place a methodological focus on the application, e.g., randomized controlled trials such as the Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and onLine TeleHealth (CONSORT-EHEALTH).

**Conclusion:** In view of the options for standards in the reporting of digital applications presented in this paper, the guideline presented by the BfArM could be expanded to include additional items in order to provide healthcare providers and users with more certainty and at the same time prevent compromises in quality standards.

### Keywords

Digital Health, Health Apps, eHealth, mHealth, Evaluation

### Sven Kernebeck MSc PH

ist seit 2019 wissenschaftlicher Mitarbeiter an der Universität Witten/Herdecke am Lehrstuhl für Didaktik und Bildungsforschung. Zuvor war er am Deutschen Zentrum für Neurodegenerative Erkrankungen, an der FH Münster am Fachbereich Gesundheit sowie an der Paracelsus Medizinische Universität (PMU) Salzburg beschäftigt. Sein Schwerpunkt: partizipative Entwicklung und die Evaluation von digitalen Technologien im Gesundheitswesen. Seit 2017 ist er Mitglied AG der Digital Health des DNVF. Kontakt: [sven.kernebeck@uni-wh.de](mailto:sven.kernebeck@uni-wh.de)



### Dr. rer. nat. Ursula Kramer

ist Geschäftsführerin der sanawork Gesundheitskommunikation. Sie engagiert sich seit 2011 als Gründerin von HealthOn für die Verbesserungen der Transparenz und Qualität digitaler Gesundheitsanwendungen. Die promovierte Apothekerin leitet zusammen mit Prof. Vollmar seit 2016 die Arbeitsgruppe Digital Health im Deutschen Netzwerk Versorgungsforschung und hat in dieser Funktion an der Entwicklung zahlreicher Publikationen mitgewirkt. Kontakt: [ursula.kramer@sanawork.de](mailto:ursula.kramer@sanawork.de)



### Dr. med. Marcus Redaelli

ist wissenschaftlicher Mitarbeiter im Institut für Gesundheitsökonomie und Klinische Epidemiologie der Universitätsklinik Köln. Seit fast zwei Jahrzehnten, u.a. an der Universität Witten/Herdecke und dem Institut für Allgemein- und Familienmedizin an der Universitätsklinik Düsseldorf, beschäftigt er sich im Rahmen der Versorgungsforschung mit telemedizinischen Möglichkeiten zur Ergänzung der Versorgungsstrukturen. Seit 2017 ist er Mitglied AG der Digital Health des DNVF. Kontakt: [marcus.redaelli@uk-koeln.de](mailto:marcus.redaelli@uk-koeln.de)



### Univ.-Prof. Dr. med. Horst Christian Vollmar MPH

hat die W3-Professur für Allgemeinmedizin an der Ruhr-Universität Bochum inne und leitet zudem die Abteilung für Allgemeinmedizin. Er ist seit 2016 Sprecher der AG Digital Health im Deutschen Netzwerk für Versorgungsforschung (DNVF).

Kontakt: [horst.vollmar@ruhr-uni-bochum.de](mailto:horst.vollmar@ruhr-uni-bochum.de)

