



Nachweis positiver Versorgungseffekte von digitalen Gesundheitsanwendungen – methodische Herausforderungen und Lösungsansätze

Inke R. König¹ · Mirja Mittermaier^{2,3} · Christian Sina⁴ · Matthias Raspe² · Patrick Stais⁵ · Thomas Gamstätter⁶ · Philipp Stachwitz⁶ · Sebastian Wolfrum⁷ · Jutta G. Richter⁸ · Martin Möckel⁹ für Arbeitsgruppe Digitale Gesundheitsanwendungen (DiGA) und künstliche Intelligenz (KI) in Leitlinien der Kommission Digitale Transformation der Deutschen Gesellschaft für Innere Medizin (DGIM)

¹ Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Universitätsklinikum Schleswig-Holstein, Campus Lübeck, Lübeck, Deutschland; ² Medizinische Klinik m.S. Infektiologie und Pneumologie, Charité – Universitätsmedizin Berlin, Berlin, Deutschland; ³ Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Deutschland; ⁴ Institut für Ernährungsmedizin, Universitätsklinikum Schleswig-Holstein, Campus Lübeck, Lübeck, Deutschland; ⁵ Medizinische Klinik 3, Lungen- und Bronchialheilkunde, Lungen-Thorax-Zentrum Nordrhein, Moers, Deutschland; ⁶ Deutsche Gesellschaft für Innere Medizin e. V., Wiesbaden, Deutschland; ⁷ Interdisziplinäre Notaufnahme, Universitätsklinikum Schleswig-Holstein, Campus Lübeck, Lübeck, Deutschland; ⁸ Poliklinik und Funktionsbereich Rheumatologie & Hiller-Forschungszentrum Rheumatologie, Medizinische Fakultät, Heinrich-Heine-Universität Düsseldorf, Universitätsklinikum Düsseldorf, Düsseldorf, Deutschland; ⁹ Notfall- und Akutmedizin, Zentrale Notaufnahmen und Chest Pain Units, Campus Mitte und Virchow-Klinikum, Charité – Universitätsmedizin Berlin, Berlin, Deutschland

Zusammenfassung

Seit 2020 sind digitale Gesundheitsanwendungen (DiGA) nach Durchlaufen eines Zulassungsverfahrens durch das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) zulasten der gesetzlichen Krankenversicherung (GKV) verordnungsfähig. Dabei können DiGA vorläufig für ein Jahr (mit der Option der Verlängerung) oder dauerhaft zugelassen werden. Letzteres ist vom wissenschaftlichen Nachweis eines positiven Versorgungseffekts abhängig, der ein medizinischer Nutzen oder eine patientenrelevante Struktur- und Verfahrensverbesserung in der Versorgung sein kann. Es zeigt sich jedoch, dass die Untersuchung von DiGA in wissenschaftlichen Studien herausfordernd ist, da es sich häufig um komplexe Interventionen handelt, in deren Erfolg auch Faktoren der Anwender und Verordner einfließen. Darüber hinaus fehlen bislang Versorgungsforschungsdaten, die den Nutzen der DiGA unterlegen. Im aktuellen Beitrag werden methodische Überlegungen zur DiGA-Forschung dargelegt. Anhand einer Auswahl internistischer DiGA wird die bisherige Forschungspraxis kritisch diskutiert.

Schlüsselwörter

Digitale Gesundheitsanwendungen/Innere Medizin · Medizinischer Nutzen · Patientenrelevante Struktur- und Verfahrensverbesserung · Studiendesign · Versorgungsforschung



QR-Code scannen & Beitrag online lesen

Seit 2020 sind digitale Gesundheitsanwendungen (DiGA) nach Durchlaufen eines Zulassungsverfahrens durch das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) zulasten der gesetzlichen Krankenversicherung (GKV) verordnungsfähig

[8]. Dabei können DiGA vorläufig [6] für ein Jahr oder dauerhaft zugelassen werden. Letzteres ist vom wissenschaftlichen Nachweis eines positiven Versorgungseffekts abhängig. Vor dem Hintergrund, dass DiGA einerseits als grundsätzlich innova-

Tab. 1 Auszug von Themen aus dem CONSORT-Statement 2010 [9, 14] in der Übersetzung von Pittler et al. [11]

Thema	Beschreibung
Hintergrund und Ziele	„Wissenschaftlicher Hintergrund und Begründung der Studie; genaue Fragestellung/en bzw. Hypothesen“
Probanden/Patienten	„Eignungskriterien der Probanden/Patienten“
Intervention/Behandlung	„Durchgeführte Intervention in jeder Gruppe mit präzisen Details, einschließlich wie und wann die Interventionen durchgeführt wurden, um Replikationen der Studie zu ermöglichen“
Endpunkte	„Vollständig definierte, primäre und sekundäre Endpunkte ..., einschließlich wie und wann sie erhoben wurden“
Verblindung	„Falls durchgeführt, wer war bei der Interventionszuordnung verblindet? (z. B. Teilnehmer, Ärzte, Therapeuten, diejenigen, die die Endpunkte beurteilten)“
Ein- und Ausschlüsse	„Für jede Gruppe Anzahl der Studienteilnehmer, die randomisiert zugeteilt wurden, die die geplante Intervention erhielten und die hinsichtlich des primären Endpunkts analysiert wurden; für jede Gruppe Zahl der Studienaussteiger und Ausschlüsse nach Randomisierung mit Angabe von Gründen“
Adhärenz und Protokollabweichungen ^a	„Definition der Adhärenz und wie diese erfasst wurde, einschließlich Ausmaß der Exposition ...; Definition von Protokollabweichungen“
Fehlende Daten ^a	„Bericht, Annahmen und statistische Methoden zum Umgang mit fehlenden Daten (z. B. multiple Imputation)“
Schaden	„Alle wichtigen Schäden ... innerhalb jeder Gruppe ...“

^aZusätzliche statistische Themen [5]
 CONSORT Consolidated Standards of Reporting Trials

tiv gelten und andererseits bei laut Gesetz niedriger Risikoklasse die unerwünschten Wirkungen als gering angenommen werden, sollten die Hürden der Zulassung nicht zu hoch sein. Es zeigt sich jedoch, dass die Untersuchung von DiGA in wissenschaftlichen Studien herausfordernd ist, da es sich häufig um komplexe Interventionen handelt, in deren Erfolg auch Faktoren der Anwender und Verordner einfließen.

Allgemeine Gesetzesgrundlagen und Leitfaden des Bundesinstituts für Arzneimittel und Medizinprodukte

Im Jahr 2019 trat das Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz [DVG]) in Kraft, das den Anspruch von Versicherten auf die Versorgung mit DiGA regelt [3]. Hierbei werden zunächst in § 33a Abs. 1 Satz 1 Sozialgesetzbuch (SGB) V die DiGA definiert. Außerdem wird hier festgelegt, dass ein Anspruch nur auf solche DiGA besteht, die „vom Bundesinstitut für Arzneimittel und Medizinprodukte in das Verzeichnis für digitale Gesundheitsanwendungen nach § 139e aufgenommen wurden ...“. In § 139e wird näher ausgeführt, dass Hersteller die Aufnahme beim BfArM beantragen und hierbei unter anderem Nachweise dazu erbringen müssen, dass die DiGA *positive Versorgungseffekte* aufweist. Dies wird dabei definiert als „ein medizinischer Nutzen oder eine pati-

entenrelevante Struktur- und Verfahrensverbesserung in der Versorgung“.

» Studien zum Nachweis positiver Versorgungseffekte müssen vergleichende Studien sein

Zur weiteren Erläuterung trat 2020 die Verordnung über das Verfahren und die Anforderungen zur Prüfung der Erstattungsfähigkeit digitaler Gesundheitsanwendungen in der gesetzlichen Krankenversicherung (Digitale-Gesundheitsanwendungen-Verordnung [DiGAV]) in Kraft [4]. Diese beschreibt in § 8 etwas detaillierter die positiven Versorgungseffekte. Zudem wird in § 10 festgelegt, dass Studien zum Nachweis positiver Versorgungseffekte vergleichende Studien sein müssen, die zeigen, dass die Anwendung der DiGA besser ist als deren Nichtanwendung. Dabei sind prospektive, aber auch retrospektive Studien mit Kontrollgruppe und im intraindividuellen Vergleich als Möglichkeiten vorgesehen.

Für konkrete Studienplanungen kann schließlich der Leitfaden des BfArM *Das Fast-Track-Verfahren für DiGA nach § 139e SGB V* herangezogen werden [2]. Hier findet sich eine tabellarische Aufstellung zu möglichen Studientypen, aus der hervorgeht, dass im Wesentlichen drei Studientypen zulässig sind. Dies wäre erstens die *beobachtende analytische Studie*, die prospektiv oder retrospektiv durchgeführt werden kann. Sie hat den Vorteil, dass

die Datenerhebung weniger aufwendig ist und dass möglicherweise auf bereits verfügbare Daten zurückgegriffen werden kann. Allerdings ist es besonders schwierig, Daten mit ausreichender Qualität zu erhalten und eine wirkliche Vergleichbarkeit der Gruppen herzustellen, was nicht nur die Zusammensetzung, sondern auch den jeweiligen Versorgungskontext betrifft. Der zweite mögliche Studientyp wäre eine *Metaanalyse*, wofür es bei den meisten aktuell entwickelten DiGA sicherlich noch zu früh ist, da das Vorhandensein valider Studien Voraussetzung wäre. Am vielversprechendsten erscheint es daher, sich auf den dritten Studientyp zu fokussieren, nämlich die *prospektive, idealerweise randomisierte, kontrollierte Interventionsstudie*. Dieser Studientyp ist am aussagekräftigsten, erfordert aber einen hohen Aufwand, zudem dauert es länger, bis Ergebnisse vorliegen.

Allgemeine Designaspekte randomisierter, kontrollierter Studien

Die allgemeinen Designaspekte für randomisierte, kontrollierte Studien (RCT), die beispielsweise im CONSORT-Statement aufgeführt werden [9, 14], sind in **Tab. 1** zusammengefasst.

Designaspekte randomisierter, kontrollierter Studien im Hinblick auf digitale Gesundheitsanwendungen

Hintergrund und Ziele

Wie oben dargelegt, fordert die Rechtsgrundlage den Nachweis, dass die Anwendung der DiGA „besser ist“ als ihre Nichtanwendung. Damit gilt allgemein, dass die DiGA selbst zu prüfen ist, nicht aber die darüber angebotene Intervention. Allerdings wird eine Reihe verschiedener Möglichkeiten für die Nichtanwendung zugelassen, denn § 10 Abs. 4 DiGAV spezifiziert:

„Die Nichtanwendung... kann eine Nichtbehandlung oder eine Behandlung ohne digitale Gesundheitsanwendung sein. Die Auswahl des Komparators muss der Versorgungsrealität entsprechen. Abweichend ... kann die Nichtanwendung auch eine Behandlung mit einer anderen vergleichbaren digitalen Gesundheitsanwendung sein.“

Damit ist offensichtlich, dass die möglichen Komparatoren sehr verschieden sein können. Das genaue Ziel der Prüfung ist daher präzise anzugeben. Entsprechend resultieren unterschiedliche Interpretationen, worauf weiter unten eingegangen wird.

Probanden/Patienten

Über die Ein- und Ausschlusskriterien der Studie wird die Zielgruppe derjenigen definiert, bei denen der positive Versorgungseffekt nachgewiesen werden soll. Bei DiGA-Studien gilt häufig als Einschlusskriterium, dass die möglichen Teilnehmenden im Sinne des Bring-your-own-device (BYOD)-Konzepts bereits Internet- oder Smartphone-Nutzende sind. Damit sind die Ergebnisse der Studie naturgemäß auf diese Personengruppe eingeschränkt und nicht darüber hinaus generalisierbar, entsprechen aber auch dem antizipierten späteren Anwenderkreis.

Ein hilfreiches Beispiel hierfür ist die vorvida-Studie von Zill et al. [16], in der die Wirksamkeit einer Internetintervention zur Reduktion des Alkoholkonsums bei Erwachsenen untersucht wurde. Die Autoren

schreiben in der Arbeit bezüglich der Generalisierbarkeit ihrer Ergebnisse: „Die untersuchte Stichprobe zeichnete sich durch einen relativ hohen Bildungs- und Beschäftigungsstatus aus. Die Übertragbarkeit der hier erhobenen Ergebnisse ist daher vor allem auf Personen bezogen, die gut im Beruf funktionieren, Zugang zum Internet haben und den Willen besitzen, sich mit einem Internet-Therapieprogramm helfen zu lassen.“

» Ein Einschlusskriterium ist häufig, dass die Teilnehmenden schon Internet oder Smartphone nutzen

Diese Einschränkung wird sich aktuell auf viele DiGA und die dazugehörigen Studien übertragen lassen, weshalb empfohlen wird, dass zum einen entsprechende Charakteristika der Stichprobe standardisiert erfasst und berichtet werden und dass auf die möglichen Limitationen hingewiesen wird. Zum anderen ist es sinnvoll, darauf hinzuwirken, dass eine DiGA einer möglichst allgemeinen Gruppe zugänglich ist, sodass sich die Zielgruppe auf patientenrelevante Weise erweitert.

Intervention/Behandlung

Die Interventionen und die DiGA selbst müssen nachvollziehbar beschrieben werden. Relevant ist zusätzlich, dass die Kontrollgruppe präzise spezifiziert wird, was im Zusammenhang mit den oben beschriebenen Zielen der Studie steht. Noch ausführlicher als in der DiGAV wird im Leitfaden des BfArM für die Vergleichsgruppe aufgeführt, dass deren Versorgung

- eine Behandlung ohne Anwendung einer DiGA,
- eine Nichtbehandlung oder
- eine Behandlung mit einer anderen, vergleichbaren DiGA

umfassen kann. Auch ein Vergleich mit der Standardbehandlung wird hier als möglich angesehen. Dies eröffnet viele Möglichkeiten, aber es muss dabei beachtet werden, dass, je nach Kontrollgruppe, unterschiedliche Effekte geprüft werden. Wenn also ein Endpunkt zwischen der Interventionsgruppe mit einer spezifischen DiGA und einer Kontrollgruppe verglichen wird, ist

vorstellbar, dass dieser Vergleich durch die folgenden Faktoren beeinflusst wird:

1. Modus der Erhebung des Endpunkts (bestenfalls digital, also im Rahmen der DiGA, vs. analog): Beispielsweise kann das Antwortverhalten bei einem Symptomfragebogen unterschiedlich sein, wenn dieser Fragebogen innerhalb der DiGA abgefragt wird anstelle einer analogen papierbasierten Version [12, 13].
2. Spezifische Intervention selbst: Beispielsweise wird in vorvida eine internetbasierte Intervention angeboten, die auf Prinzipien und Techniken der kognitiven Verhaltenstherapie aufbaut. Hierbei ist zu vermuten, dass die vier Therapiemodule selbst einen Effekt haben, der unabhängig davon ist, ob sie in dieser speziellen DiGA internetbasiert angeboten werden.
3. Digitalisierte Form der Intervention, also Umsetzung in der DiGA: Zusätzlich zur eigentlichen Intervention kann die digitale Umsetzung einen eigenen Effekt haben, weil sie beispielsweise eine größere Personalisierung und Flexibilität ermöglicht und gleichzeitig unabhängig vom spezifischen Behandelnden ist.
4. Spezifische Form der Digitalisierung, also spezielle DiGA: Die Implementierung der DiGA kann einen zusätzlichen Effekt haben dadurch, dass die DiGA besondere Usability-Aspekte aufweist (etwa bezüglich Handhabbarkeit, „graphical user interface“).

Es ist plausibel, dass sich der gemessene Effekt, also der Unterschied zur Interventionsgruppe, dann aus den genannten Effekten zusammensetzt. Unter der Annahme, dass der Endpunkt im selben Modus erhoben wird (Punkt 1), würden bei verschiedenen Kontrollgruppen dann die Effekte wie in **Tab. 2** eine Rolle spielen.

Bei den meisten publizierten DiGA-Studien erhält die Kontrollgruppe eine nicht-digitale Standardbehandlung. Damit wäre der gemessene Effekt letztendlich am stärksten, aber auch nicht unbedingt nur auf die spezielle DiGA zurückzuführen. Es wird daher empfohlen, dies präzise zu berichten und zu diskutieren. Mit der erwarteten Zunahme an verfügbaren DiGA ist aber auch absehbar, dass mehr Vergleiche

Tab. 2 Zusammensetzung des Effekts beim Vergleich der Interventionsgruppe mit verschiedenen möglichen Kontrollgruppen

Kontrollgruppe	Effekt ...		
	... der Intervention	... der Digitalisierung	... der spezifischen DiGA
Nichtdigitale Standardbehandlung	x	x	x
Wartekontrolle ^a	x	x	x
Nichtdigitale Intervention	–	x	x
Anderer DiGA mit anderer Intervention	x	–	x
Anderer DiGA mit selber Intervention	–	–	x

^aPatienten, die auf die Verordnung der DiGA warten
DiGA digitale Gesundheitsanwendung

che mit bereits gelisteten DiGA als Kontrollgruppe möglich sein werden, sodass die gemessenen Effekte geschärft werden können.

Endpunkte

Nach § 8 Abs. 2 DiGAV ist der medizinische Nutzen, also der patientenrelevante Effekt, nachzuweisen, insbesondere hinsichtlich

- der Verbesserung des Gesundheitszustands,
- der Verkürzung der Krankheitsdauer,
- der Verlängerung des Überlebens,
- einer höheren Remissionsrate oder
- einer Verbesserung der Lebensqualität.

Unabhängig vom spezifischen Endpunkt ist hierbei die Beobachtungsgleichheit wichtig, was wiederum (siehe Abschnitt „Intervention/Behandlung“) dafürspricht, den Modus der Erhebung über die Gruppen hinweg gleich zu halten – bei den meisten DiGA-Studien erfolgt dies digital.

» Meist werden subjektiv erhobene, patientenzentrierte Endpunkte eingesetzt

Meist werden wie in anderen Studien zu digitalen Applikationen subjektiv erhobene, patientenzentrierte Endpunkte eingesetzt, sodass die Teilnehmenden selbst ihren Gesundheitszustand, ihre Krankheitsaktivität oder die Lebensqualität mithilfe von Patient-reported-outcome(PRO)-Instrumenten einschätzen [12, 13]. Der Vorteil ist, dass dies einfach digital umsetzbar ist und somit, zumindest für die Interventionsgruppe, einen nur geringen Mehraufwand bedeutet. Außerdem können patientenzentrierte digital erfasste Endpunkte

häufiger gemessen werden, sodass insgesamt die Erhebung von Verlaufsdaten einfacher wird. Dies kann im Umgang mit fehlenden Werten hilfreich sein (siehe Abschnitt „Fehlende Daten“).

Von Nachteil dabei ist allerdings, dass die Qualitätskontrolle der Endpunkte schwieriger ist, weil es im engeren Sinne keine Quelldaten gibt, mit denen die erfassten Daten abgeglichen werden können. Dies ist allerdings nicht spezifisch für DiGA-Studien, sondern gilt in gleicher Weise für alle Studien, bei denen die Teilnehmenden selbstständig Endpunkte erfassen. Ein weiterer Nachteil besteht darin, dass mögliche Verzerrungen im Falle einer fehlenden Verblindung (siehe Abschnitt „Verblindung“) bei objektiv gemessenen Endpunkten geringer ausfallen. Es ist daher zu empfehlen, zusätzlich (sekundäre) Endpunkte zu verwenden, die leicht zu objektivieren sind. Beispielsweise werden in der Studie zur Prüfung der DiGA zanadio zur Behandlung von Menschen mit Adipositas die objektivierbaren Parameter prozentuale Gewichtsreduktion und „waist-to-hip ratio“ gemessen [10].

Verblindung

Wenn der Kontrollgruppe nicht gerade eine andere DiGA zur Verfügung gestellt wird, ist eine Verblindung der Teilnehmenden in der Regel nicht möglich. Von einer Verblindung der Behandelnden zu sprechen, ist bei einer DiGA nicht zutreffend, wenn die Intervention selbst ausschließlich über die DiGA, also ohne menschliche Interaktion, zur Verfügung gestellt wird. Schließlich bleibt die prinzipielle Möglichkeit, diejenigen zu verblenden, die das Zielkriterium erheben. Allerdings ist dies na-

turgemäß nicht der Fall, wenn patientenzentrierte Zielkriterien gemessen werden, die von den Teilnehmenden selbst angegeben werden. Entsprechend werden DiGA-Studien typischerweise als offene Studien durchgeführt.

Damit ist es schwieriger, den Effekt der DiGA von anderen Effekten wie einem Placeboeffekt zu unterscheiden. Für einen Kausalitätsnachweis sind dann möglicherweise zusätzliche Plausibilisierungen hilfreich. Beispielsweise könnte in vielen Fällen das Ausmaß der Exposition, also die „Dosis“ der DiGA bzw. die Adhärenz zur DiGA-Nutzung, mit dem Effekt korreliert werden.

Einschlüsse und Abbruchquoten

Wenn im Lauf der Studie zu viele Teilnehmende die Teilnahme an der Studie abbrechen, kann die Strukturgleichheit der Randomisierung und auch die Repräsentativität gefährdet werden. Außerdem wird durch eine zu geringe Fallzahl die Power der Studie verringert. Damit ist es wichtig, schon bei der Planung der Studie Risiken für hohe Abbruchraten zu erkennen und möglichst gering zu halten. Bei DiGA-Studien könnten die Abbruchraten höher als in herkömmlichen klinischen Studien sein, weil es nur wenig direkte menschliche Interaktion während der Intervention gibt. Dies kann dadurch verkompliziert werden, dass die Abbruchraten sich in den Therapiearmen auf nicht zufällige Weise unterscheiden, weil beispielsweise der Aufwand in der Interventionsgruppe als zu hoch wahrgenommen wird. So berichten beispielsweise Zill et al. [16] zur vorvida-Studie eine Abbruchquote von 32% in der Interventionsgruppe, aber nur von 17% in der Kontrollgruppe ohne eine weitere Angabe von Gründen.

Weil der allgemeine positive Versorgungseffekt der DiGA logischerweise die Nutzung der DiGA voraussetzt, könnte in einigen Fällen darüber nachgedacht werden, auch den Grad der Nutzung selbst als zusätzlichen Endpunkt neben der eigentlichen Wirksamkeit zu berücksichtigen. Damit wäre der Effekt nur dann nachgewiesen, wenn erstens die DiGA in einem im Studienprotokoll definierten Ausmaß genutzt und zweitens ein patientenrelevanter Unterschied nachgewiesen wurde. Hier

sollte auch dem Umstand Rechnung getragen werden, dass das Ausmaß notwendiger Nutzung in Abhängigkeit von der jeweiligen DiGA unterschiedlich ausfallen kann.

Adhärenz und Protokollabweichungen

Auch bei der Definition von Protokolladhärenz und Analysepopulationen ist wichtig, dass sich dies auf die Nutzung der DiGA, aber nicht unbedingt auf die Umsetzung der Intervention bezieht.

Eine besonders ausführliche Betrachtung dieser Definitionen findet sich in der Studie *deprexis*, die die Effekte einer internetbasierten Intervention auf Symptome schwerer Depression untersucht hat [7]. Wichtig für das Verständnis der Protokolladhärenz ist, dass die Teilnehmenden randomisiert wurden und bei Randomisierung in den Interventionsarm einen DiGA-Gutschein erhielten, mit dem sie dann die internetbasierte Intervention freischalten konnten. Es wurden hier schließlich drei verschiedene Gruppen gebildet, nämlich

- die Gruppe aller randomisierten Teilnehmenden,
- die Gruppe derjenigen Teilnehmenden der Interventionsgruppe, die ihren Gutschein aktiviert und damit ihr Interesse an der Intervention dokumentiert hatten, und
- die Gruppe der Teilnehmenden, die mindestens 4 Sitzungen der Intervention begonnen und mindestens 60 min aktiv am Programm teilgenommen hatten.

Nach dieser Einteilung konnte dann im Flussdiagramm dargestellt werden, dass etwa 10% der Teilnehmenden der Interventionsgruppe den Gutschein nicht aktivierten und weitere 18% nicht aktiv am Programm teilnahmen. Der Versorgungseffekt konnte dann separat in diesen Gruppen geschätzt werden.

Um nachzuvollziehen, in welchem Ausmaß ein möglicherweise gemessener Effekt auf die Nutzung der DiGA zurückzuführen ist, ist eine solche ausführliche Darstellung hilfreich. Daher ist zu empfehlen, bei einer DiGA-Studie die tatsächliche Nutzung der DiGA und der darüber angebotenen Intervention zu messen. Dazu

müssen ethische und datenschutzrechtliche Vorgaben erfüllt sein und die Zustimmungen der Teilnehmenden vorliegen. Es ist allerdings wichtig zu beachten, dass je nach DiGA die Adhärenz unterschiedliche Bedeutung haben kann. Wird beispielsweise eine Intervention per DiGA erfolgreich vermittelt, kann es sein, dass die Patienten die Nutzung reduzieren oder einstellen, der Interventionserfolg aber dennoch vorhanden ist.

Fehlende Daten

Aus den obigen Abschnitten ging bereits hervor, dass durch Ausschlüsse bzw. Abbrüche in verschiedenen Phasen der Studie fehlende Werte entstehen. Zusätzlich kommt es vor, dass aktiv Teilnehmende Fragen nicht beantworten oder aus weiteren Gründen Werte nicht vorliegen. Für den Umgang mit fehlenden Werten gibt es verschiedene Möglichkeiten, die von Ausmaß und Natur der fehlenden Werte abhängen. Hierfür werden speziell Situationen unterschieden, in denen die Wahrscheinlichkeit dafür, dass ein Wert fehlt, vom Wert selbst oder vom Wert in anderen Variablen abhängt oder von allen unabhängig ist. Wenn beispielsweise präferenziell Personen die Teilnahme abbrechen, denen es im Hinblick auf den Endpunkt besonders schlecht geht, wäre das Fehlen des Endpunktwerts eben von genau diesem Wert abhängig. Wenn hingegen Personen der Interventionsgruppe die Teilnahme abbrechen, denen die Handhabung der DiGA zu kompliziert erscheint, könnte dies mit dem Alter oder Bildungsgrad zusammenhängen. In diesem Fall wäre das Fehlen zusätzlich vom Therapiearm abhängig.

Im Rahmen von DiGA-Studien ist bislang noch wenig dazu bekannt, welche Mechanismen dazu führen, dass Teilnehmende Fragen nicht beantworten, die DiGA nicht nutzen oder die Teilnahme insgesamt abbrechen. Weil dies auch für den Umgang mit den daraus entstehenden fehlenden Werten wichtig ist, sind eine Aufschlüsselung der Gruppen (siehe Abschnitt „Ein- und Ausschlüsse“), eine Differenzierung der DiGA-Adhärenz (siehe Abschnitt „Adhärenz und Protokollabweichungen“) und eine Erfassung der jeweiligen individuellen Gründe besonders nützlich.

In herkömmlichen klinischen Studien ist es inzwischen Standard, fehlende Werte durch multiple Imputation zu ersetzen, wofür nichtfehlende Werte der Teilnehmenden genutzt werden. Wie oben beschrieben, kann es in DiGA-Studien einfacher sein, Verlaufsdaten zu erheben, und dies kann ein entscheidender Vorteil für den Einsatz von Imputationstechniken sein.

Schaden

Der Bericht von Schäden und unerwünschten Ereignissen spielt bei herkömmlichen klinischen Studien eine große Rolle. Obwohl alle Ereignisse unabhängig von ihrer vermuteten Kausalität zu erfassen sind, ist im Speziellen bei DiGA-Studien die Sicherheit bezogen auf die DiGA selbst gemeint, nicht auf die darüber dargebotene Intervention. In der Regel ist hier kein Schaden zu erwarten und so wird in den meisten Publikationen von DiGA-Studien auch nichts dazu berichtet. Ob dies wirklich zutrifft, bleibt allerdings abzuwarten, und alle Studien sollten potenzielle unerwünschte Wirkungen erfassen.

Für die Anwendung einer Reihe von DiGA existieren Kontraindikationen, die jeweils im DiGA-Verzeichnis des BfArM aufgeführt werden. So werden beispielsweise für die DiGA *VIVIRA*, die auf Schmerzreduktion durch Bewegungstherapie bei Rückenschmerzen abzielt, neben 45 Indikationen auch 27 Kontraindikationen genannt. Hierzu zählen unter anderem knöchernen Tumorinfiltrationen oder Nervenschäden bei Bandscheibenvorfällen. Auch bei DiGA zur Behandlung von Angst- und Panikstörungen finden sich jeweils Kontraindikationen wie Psychosen oder Abhängigkeitssyndrome.

» In der Regelversorgung ist denkbar, dass Patienten eine DiGA trotz Kontraindikationen anwenden

Unter Studienbedingungen kann verhindert werden, dass Patienten mit Kontraindikationen eine DiGA nutzen. Unter den Bedingungen der Regelversorgung sowie der aktuellen Zugangswege zu DiGA und mit steigenden Nutzungszahlen ist es durchaus denkbar, dass Patienten eine DiGA trotz gegebener Kontraindikationen anwenden. Patienten können auch ohne

ärztliche Verordnung die Nutzung einer DiGA allein durch Nachweis einer entsprechenden Diagnose von ihrer Krankenkasse erstatten lassen, die jedoch in der Regel das Vorliegen von Kontraindikationen nicht prüfen kann.

Diskussion

Zum Zeitpunkt der Manuskripteinreichung (15.04.2022) sind 5 von 32 DiGA für Patienten mit einer typisch internistischen Indikation (Diagnose) verfügbar:

- Cara Care für Reizdarm
- ESYSTA-App und -Portal für digitales Diabetesmanagement
- HelloBetter Diabetes und Depression
- Oviva Direkt für Adipositas
- zanadio für Adipositas

Die DiGA ViViRA lässt sich sowohl der Rheumatologie als auch der Orthopädie zuordnen, und die DiGA NichtraucherHelden adressiert ein vordringliches Thema der Prävention. Nachfolgend soll das methodische Studiendesign beispielhaft an 5 der internistisch relevanten DiGA diskutiert werden. Die meisten dieser DiGA sind bisher vorläufig zugelassen, nur ViViRA seit dem 17.02.2022 dauerhaft. Dies (nach einer zweimal verlängerten Erprobungsphase) allerdings nur für die Indikationsgruppe Rückenschmerzen, da der Nachweis positiver Versorgungseffekte für die ursprünglich mit adressierten Indikationsbereiche Knie- und Hüftschmerz nicht erbracht werden konnte.

Zanadio/Oviva Direkt für Adipositas

Die DiGA zanadio ist vorläufig gelistet, und es liegen noch keine finalen Studienergebnisse vor. Das Studiendesign ist sowohl auf der Seite des BfArM (<https://diga.bfarm.de/de/verzeichnis/1346/fachkreise>; Zugriff: 30.01.2022) als auch auf der Website des Deutschen Registers Klinischer Studien (DRKS; www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00024415; Zugriff: 23.01.2022) beschrieben, jedoch unvollständig und zum Teil widersprüchlich. Die 2-armige experimentelle Interventionsstudie mit 150 adipösen Patienten (77 in der Interventionsgruppe – Nutzung der DiGA, 73 in der Kontrollgruppe – keine Inter-

vention) hat das Ziel, die Wirksamkeit der DiGA (positiver Versorgungseffekt) zu evaluieren [10]. Primärer Endpunkt ist eine Gewichtsreduktion um 5% in 12 Monaten, sekundäre Endpunkte sind die „waist-to-hip ratio“ sowie Wohlbefinden und Lebensqualität. Da die Kontrollgruppe keine Behandlung erhält, bleibt unklar, ob ein positiver Versorgungseffekt auf den Inhalt der DiGA, auf die Nutzung einer digitalen Anwendung oder lediglich auf einen Placeboeffekt zurückzuführen ist (■ Tab. 2).

Neben zanadio ist auch Oviva Direkt für Adipositas für die Indikation E66 Adipositas im DiGA-Verzeichnis aufgeführt. Es gibt bisher keine Richtlinien oder Empfehlungen, wie ein Verordner vorgehen soll, wenn 2 oder mehrere DiGA für dasselbe Indikationsgebiet verfügbar sind. Eine Möglichkeit für zukünftige Studien wäre, eine bereits bestehende DiGA als Kontrollintervention zu nutzen.

Cara Care

Die DiGA Cara Care für Reizdarm dient der Eigenanwendung zur digitalen Behandlung des Reizdarmsyndroms (RDS). Die DiGA Cara Care ist vorläufig im DiGA-Verzeichnis gelistet. Finale Studienergebnisse liegen noch nicht vor. Das Studiendesign ist sowohl auf der Seite des BfArM (<https://diga.bfarm.de/verzeichnis/294/fachkreise>; Zugriff: 23.01.2022) als auch auf der Website des DRKS (https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00026631; Zugriff: 23.01.2022) beschrieben. Laut BfArM-Verzeichnis handelt es sich um eine prospektive, multizentrische, 2-armige, randomisierte Studie mit parallelem Gruppensdesign. Die Interventionsgruppe erhält den Zugang zur Vollversion der DiGA Cara Care für Reizdarm und darf sich parallel zur DiGA-Anwendung in ärztliche Behandlung gemäß Regelversorgung begeben (Standardbehandlung). Die Kontrollgruppe bekommt keinen DiGA-Zugang (Nichtanwendung). Auch dieser Gruppe steht die Regelversorgung offen.

Der primäre Endpunkt wird nach 12 Wochen Intervention anhand des Irritable Bowel Syndrome Severity Scoring System (IBS-SSS) ermittelt. Nicht angegeben wird,

wie groß der erwartete Effekt (Veränderung des IBS-SSS) gegenüber der Nichtanwendung ist. Sekundäre Endpunkte sind unter anderem Empfinden von Angst (gemessen mittels Generalized Anxiety Disorder Scale-7 [GAD-7]), Empfinden von Depressivität (gemessen mittels Patient Health Questionnaire-9 [PHQ-9]) und Lebensqualität (gemessen mittels Irritable Bowel Syndrome Quality of Life [IBS-QOL]). Die geplante Studienteilnehmerzahl beträgt 374.

Wie bereits für die Angaben zur zanadio-Studie muss festgestellt werden, dass sowohl das BfArM-Verzeichnis als auch das DRKS nur unzureichend über die Intervention informieren, sodass eine abschließende Beurteilung des Studiendesigns nicht möglich ist. Ebenfalls analog zur zanadio-Erprobung gilt es festzuhalten, dass die Kontrollgruppe vermutlich keine Behandlung erhält; diesbezüglich ist die Angabe des DRKS nicht eindeutig.

Wichtig erscheint, dass die Standardbehandlung – zumindest aus den vorhandenen Informationen – nicht klar definiert wird. Gerade für das RDS sind diese Angaben von Bedeutung, da die DiGA zugrunde liegenden Wirkprinzipien (niedrige fermentierbare Oligosaccharide, Disaccharide, Monosaccharide und Polyole [FODMAP] und mit Einschränkungen auch organzentrierte Entspannungsübungen/„Darmhypnose“) laut aktueller S3-Leitlinie (<https://www.awmf.org/leitlinien/detail/II/021-016.html>) Teil der RDS-Therapie und damit Standardbehandlung sind. Die Erfahrung zeigt aber, dass diese Therapien gerade im ländlichen Raum nicht immer im klinischen Alltag zur Verfügung stehen. Dies eröffnet ein weiteres Problemfeld der Erprobung von DiGA und sollte bei der Studienplanung bedacht werden.

ESYSTA

ESYSTA dient insulinbehandelten Diabetiker:innen zum Selbstmanagement in häuslicher Umgebung. Die DiGA wurde vorläufig zugelassen. Studienergebnisse mit der DiGA stehen aus, die Wirksamkeit wird aus allgemeinen Erkenntnissen zum Diabetesselbstmanagement abgeleitet. Das Ziel einer geplanten randomisierten Studie ist die Absenkung des Hämoglobin-

A1c(HbA1c)-Werts innerhalb von 6 Monaten. Die Studie soll beim DRKS eingetragen werden, was am 02.02.2022 noch nicht der Fall war (Suche nach „ESYSTA“ unter https://www.drks.de/drks_web/navigate.do?sessionId=84E566ED9B5297FBEC218DA32E41789?navigationId=results; Zugriff: 02.02.2022).

Die geplante Studie soll nach Angaben unter „Informationen für Fachkreise“ im DiGA-Verzeichnis [1] „mindestens“ 166 Patientinnen und Patienten mit Diabetes mellitus Typ 1 oder 2 einschließen und dann die Anwendung von ESYSTA als Teil des Diabetesmanagements im Vergleich mit den Daten einer randomisierten Kontrollgruppe, die keine Diabetes-App nutzt, evaluieren.

Als Ergebnisparameter werden Veränderungen des HbA1c-Werts erfasst. Im Sinne unserer Ausführungen und unter Bezug auf **Tab. 2** wird diese Studie bei einer Kontrollgruppe mit nichtdigitaler Standardbehandlung einen Effekt der Intervention, der Digitalisierung und der spezifischen DiGA untersuchen. Die Problematik dabei ist, dass diese Effekte nicht voneinander abgrenzbar sind und so beispielsweise Änderungen der DiGA zukünftig eine Übertragbarkeit der Studienergebnisse fraglich machen.

ViViRA

Die DiGA ViViRA, die ursprünglich auf die „Schmerzreduktion bei Patienten mit Rücken-, Knie- und Hüftschmerzen“ zielte (<https://www.vivira.com/presse/>), ist im DiGA-Verzeichnis des BfArM seit 17.02.2022 dauerhaft gelistet, allerdings nur für Rückenschmerzen (<https://diga.bfarm.de/de/verzeichnis/387/fachkreise>; Zugriff: 25.01.2022). Das DRKS listet aktuell (Stand 25.01.2022) 3 Studien mit ViViRA auf, von denen 2 als abgeschlossen gekennzeichnet sind (laut wissenschaftlicher Kurzbeschreibung des DRKS: „Prospektive randomisierte, gruppenkontrollierte Studie von Wirksamkeit und Sicherheit der Vivira-App zur Behandlung von Wirbelsäulen-/Rückenschmerzen“, gefördert durch die Bundesministerin für Forschung und Technologie c/o VDI/VDE Innovation+Technik GmbH, angestrebte Teilnehmerzahl $n=198$, https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00022781, und „Retrospektive, intraindividuell kontrollierte Studie mit Einschluss von 197 Patienten mit chronischen Schmerzen des muskuloskeletalen Systems“, https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00021785). Es stehen nach Rücksprache mit dem DiGA-Hersteller noch keine finalen, publizierten Studienergebnisse zur Verfügung (Stand 24.01.2022). Diese werden erst für 2022 erwartet. Die „Retrospektive, intraindividuell kontrollierte Studie mit Einschluss von 200 Patienten mit chronischen Schmerzen des muskuloskeletalen Systems“ (https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00024051) rekrutiert gemäß DRKS-Eintrag noch Teilnehmer.

» **Wie vergleichbare Ausschlusskriterien zukünftig gewährleistet werden können, bleibt abzuwarten**

Alle 3 Studien werden bzw. wurden monozentrisch durchgeführt. Die noch rekrutierende Studie (https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00024051; Stand 25.01.2022) listet umfangreiche Ausschlusskriterien für die Teilnahme auf, darunter die „bisherige Anwendung einer Bewegungstherapie DiGA (Vivira, Kaia, etc.)“. Das Auflisten entsprechender Ausschlusskriterien ist sicher für alle Beteiligten noch ungewohnt, wird aber in Zukunft eine Rolle bei Studien zu bzw. mit DiGA spielen. Wie vergleichbare Ausschlusskriterien zukünftig gewährleistet werden können, bleibt abzuwarten.

NichtraucherHelden

Mit der DiGA NichtraucherHelden wird eine Raucherentwöhnung unterstützt. Der betreffende Code nach Internationaler statistischer Klassifikation der Krankheiten und verwandter Gesundheitsprobleme (ICD) für die Verordnung ist „F17.2 Psychische und Verhaltensstörungen durch Tabak: Abhängigkeitssyndrom“. Die DiGA ist seit dem 03.07.2021 vorläufig im BfArM-Verzeichnis aufgenommen (<https://diga.bfarm.de/de/verzeichnis/1085>; Zugriff: 22.02.2022). NichtraucherHelden wird

initial für 3 Monate verordnet. Folgeverordnungen für jeweils weitere 3 Monate sind möglich (dann unter anderer Pharmazentralnummer [PZN] mit geringerem Preis). Eine Kontraindikation stellen vor allem instabile psychische Erkrankungen dar. In verschiedenen Modulen werden 3 Phasen – Vorbereitung, Rauchstopp und Stabilisierung – durchlaufen. Die Methodik orientiert sich eng an in aktuellen Leitlinien empfohlenen und durch breite Evidenz belegten Vorgehen eines verhaltenstherapeutischen Ansatzes. Bei der vorläufigen Zulassung wurde unter anderem eine 1-armige, nichtverblindete Pilotstudie mit 50 Proband:innen berücksichtigt. Laut Informationen im BfArM-Verzeichnis gingen die Datensätze von 40 Proband:innen in die finale Analyse ein. Nach 4 Monaten App-Anwendung lag die Raucherabstinenz bei 30% (7-Tages-Prävalenz). Die gesundheitsbezogene Lebensqualität wurde mit dem Short Form-12 (SF-12) gemessen. Im Vergleich zum Studienbeginn und zur Gruppe der Weiterrauchenden verbesserte sich die körperliche Lebensqualität unter den erfolgreichen Nichtrauchern signifikant und die psychische „leicht“.

Im DRKS wurde am 30.07.2021 eine nichtverblindete RCT angemeldet (https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00025933; Zugriff: 22.02.2022), die am 09.11.2021 erstmalig eine Proband:in einschloss. Die multizentrische Studie in pneumologischen und allgemeinärztlichen Praxen in Berlin und Baden-Württemberg plant eine Teilnehmerzahl von insgesamt 640, die in 2 Arme (3-minütige ärztliche Kurzberatung mit vs. ohne DiGA-Verschreibung; jeweils $n=320$) randomisiert werden sollen. Neben Ein- und Ausschlusskriterien werden primäre und sekundäre Endpunkte der Studie definiert. Der primäre Endpunkt ist die 7-Tages-Punktprävalenz der Rauchabstinenz 6 Monate nach dem Rauchstopp. Sekundäre Endpunkte sind eine prolongierte Rauchabstinenz, die objektive Rauchabstinenz mittels Cotininnachweis im Speichel nach 6 Monaten und Veränderungen von gesundheitsbezogener Lebensqualität, Luftnot und Husten (jeweils mit etablierten Fragebögen gemessen). Es wurden laut Protokoll 5 Evaluationszeitpunkte

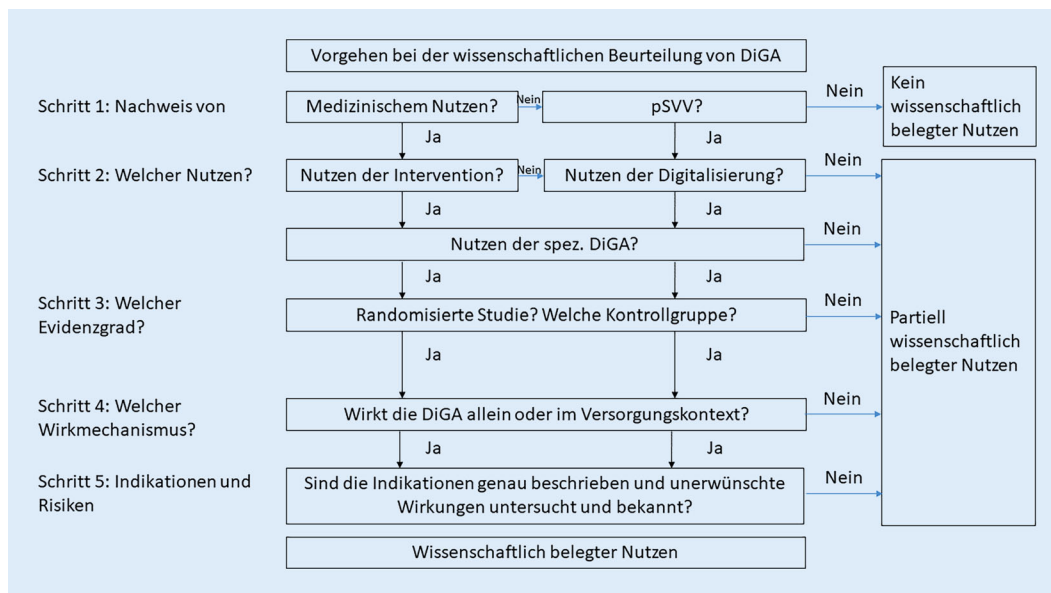


Abb. 1 ◀ Vorgehen bei der wissenschaftlichen Beurteilung einer DiGA auf der Basis publizierter Daten. DiGA digitale Gesundheitsanwendung, pSVV patientenrelevante Struktur- und Verfahrensverbesserungen

festgelegt: vor Beginn = t0, 24 h (entfällt im Kontrollarm), 4 Wochen, 3 und 6 Monate nach Beginn = t1–4. Wie der aktuelle Rekrutierungsstand ist bzw. wann Ergebnisse der Untersuchung zu erwarten sind, ist nicht bekannt.

Das Studiendesign ist vergleichsweise detailliert beschrieben. Nachteilig ist, dass nicht berücksichtigt wird, ob zeitgleich eine Nikotinersatztherapie oder Einnahme anderer Medikamente zur Linderung der Entzugssymptomatik erfolgt, was als Goldstandard in der Tabakentwöhnung gilt und höhere Erfolgsquoten mit sich bringt. Auch die Nutzung der DiGA wird nach Angaben im Studienprotokoll nicht quantifiziert. Effekte der Digitalisierung bzw. der spezifischen DiGA sind nicht klar zu trennen.

Fazit

Zusammenfassend kann festgestellt werden, dass einerseits die bisherigen wissenschaftlichen Daten zu internistischen DiGA noch unzureichend sind und andererseits geplante Studien oft nur auf Teilaspekte wünschenswerter Evidenz abzielen. Untersuchungen zur Wirksamkeit von DiGA in der Routineanwendung, beispielsweise auf Basis von Abrechnungsdaten, wurden bisher nicht publiziert. Kritik findet sich auch in einem aktuellen DiGA-Report der Techniker Krankenkasse (DiGA-Report 2022; <https://www.tk.de/resource/blob/2125136/dd3d3dbafcaef0984dcf8576b1d7713/tk-diga-report-2022-data.pdf>; Zu-

griff: 15.04.2022), in dem die folgenden Punkte herauszuheben sind:

- Erforderlich ist die Einigung auf eine etablierte Methodik zur Bewertung des „Evidenzkörpers“, vorgeschlagen werden die GRADE-Prinzipien (GRADE Grading of Recommendations, Assessment, Development and Evaluation; S. 68 f.).
- Etablierung eines „Verfahrensautomatismus“, der die langfristige (Re-)Evaluierung von DiGA anhand von Versorgungsdaten (beispielsweise „GKV-Abrechnungs- und Nutzungsdaten“) quasi erzwingt (S. 73). Dies wird auch von anderen gefordert, die die Bedeutung von Real-world-Daten betonen [15].

In **Abb. 1** wird zusammengefasst, wie pragmatisch eine DiGA unter wissenschaftlichen Aspekten beurteilt werden kann. Zugleich dient diese Abbildung als erste Orientierung bei der Studienplanung. Dabei sollten die Studien, wie oben dargelegt, sinnvolle, vorab definierte Fragestellungen untersuchen und nicht nur „irgendwelche Effekte“ signifikant nachweisen.

Korrespondenzadresse

Univ.-Prof. Dr. Martin Möckel
Notfall- und Akutmedizin, Zentrale Notaufnahmen und Chest Pain Units, Campus Mitte und Virchow-Klinikum, Charité – Universitätsmedizin Berlin
Charitéplatz 1, 10117 Berlin, Deutschland
martin.moekel@charite.de

Einhaltung ethischer Richtlinien

Interessenkonflikt. I.R. König, M. Mittermaier, C. Sina, M. Raspe, P. Stais, T. Gamstätter, P. Stachwitz, S. Wolfrum, J.G. Richter und M. Möckel geben an, dass kein Interessenkonflikt besteht.

Für diesen Beitrag wurden von den Autoren/-innen keine Studien an Menschen oder Tieren durchgeführt. Für die aufgeführten Studien gelten die jeweils dort angegebenen ethischen Richtlinien.

Literatur

1. BfArM (2022) DiGA-Verzeichnis. Bundesinstitut für Arzneimittel und Medizinprodukte. <https://diga.bfarm.de/de/verzeichnis?type=%5B%5D>. Zugegriffen: 03.10.2022
2. BfArM (2022) Leitfaden und Ausfüllhilfe. Bundesamt für Arzneimittel und Medizinprodukte. https://www.bfarm.de/DE/Medizinprodukte/Portale/DiGA/_node.html. Zugegriffen: 03.10.2022
3. Bundestag (2019) Digitale-Versorgung-Gesetz – DVG. Bundesgesetzblatt 2019:2562–2584
4. Bundestag (2020) Digitale Gesundheitsanwendungen-Verordnung – DIGAV. Bundesgesetzblatt 2020:768–798
5. Gamble C, Krishan A, Stocken D et al (2017) Guidelines for the content of statistical analysis plans in clinical trials. J Am Med Assoc 318:2337–2343
6. Heimann P, Lorenz N, Blum N et al (2021) Experiences of digital health care applications

- (DiGA) manufacturers with the BfArM Fast-Track procedure. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 64:1249–1253
7. Meyer B, Bierbrodt J, Schröder J et al (2015) Effects of an internet intervention (deprexis) on severe depression symptoms: randomized controlled trial. *Internet Interv* 2:48–59
 8. Mittermaier M, Sina C, Richter JG et al (2022) Practical use of digital health applications (DiGA) in internal medicine. *Internist* 63:245–254
 9. Moher D, Hopewell S, Schulz KF et al (2010) CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Br Med J* 340:c869
 10. Ordnung M, Horstmann A (2021) Die wissenschaftliche Evaluierung von zanadio – einem ganzheitlichen, digitalen Behandlungsprogramm für Menschen mit Adipositas. *Adipositas* 15:152–153
 11. Pittler MH, Blümle A, Meerpohl JJ et al (2011) CONSORT 2010: Aktualisierte Leitlinie für Berichte randomisierter Studien im Parallelgruppen-Design. *Dtsch Med Wochenschr* 136:e20–e23
 12. Richter JG, Nannen C, Chehab G et al (2021) Mobile App-based documentation of patient-reported outcomes—3-months results from a proof-of-concept study on modern rheumatology patient management. *Arthritis Res Ther* 23:121
 13. Richter JG, Weiß A, Bungartz C et al (2021) Mobile responsive app—A useful additional tool for data collection in the German pregnancy register rhekiss? *Front Med* 8:773836
 14. Schulz KF, Altman DG, Moher D (2010) CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med* 152:726–732
 15. Stern AD, Brönneke J, Debatin JF et al (2022) Advancing digital health applications: priorities for innovation in real-world evidence generation. *Lancet Digit Health* 4:e200–e206
 16. Zill JM, Christalle E, Meyer B et al (2019) The effectiveness of an internet intervention aimed at reducing alcohol consumption in adults. *Dtsch Arztebl Int* 116:127–133

Evidence of positive care effects by digital health apps—methodological challenges and approaches

Since 2020, digital health applications (DiGA) can be prescribed at the expense of the German statutory health insurance (SHI) system after undergoing an approval procedure by the Federal Institute for Drugs and Medical Devices (BfArM). DiGA can be approved provisionally for 1 year (with the option of extension) or permanently. The latter is dependent on scientific evidence of a positive effect on care, which can be a medical benefit or a patient-relevant structural and procedural improvement in care. However, it is apparent that the investigation of DiGA in scientific studies is challenging, as they are often complex interventions whose success also includes user and prescriber factors. In addition, health services research data underpinning the benefits of DiGA are lacking to date. In the current article, methodological considerations for DiGA research are presented, and a selection of internal medicine DiGAs is used to critically discuss current research practice.

Keywords

Digital health applications/internal medicine · Medical benefit · Patient-relevant structural and procedural improvement · Study design · Health services research