



Medizinische Begutachtung digitaler Gesundheitsanwendungen im Projekt „Trusted-Health-Apps“

Methodendokumentation

Medizinische Begutachtung digitaler Gesundheitsanwendungen im Projekt „Trusted-Health-Apps“

Methodendokumentation

Version: 1.0.0

Autorinnen und Autoren

Sebastian Kuhn, Timo Thranberend

Sina Busse, Stefanie Hennig, Maike Henningsen,
Stefanie Jungmann, Jochen Klucken, Johanna Ludwig,
Sabine Maur, Martin Möckel, Olaf Schmidt,
Christoph Schöbel, Robin Weidemann

Inhalt

1	Einleitung und Zusammenfassung	6
2	Projekthintergrund	8
3	Definitionen	10
3.1	Digitale Gesundheitsanwendungen (DiGA)	10
3.2	AppQ	11
3.3	Medizinisch-fachliche Fundierung	12
3.4	Qualität der Evidenz der positiven Versorgungseffekte	12
4	Herausforderung: Medizinische Bewertung von DiGA	13
5	Der Begutachtungsprozess	14
5.1	Zielsetzung	14
5.2	Bewertungsgegenstand	14
5.3	Ausgangsbasis	15
5.4	Rollen und Verantwortlichkeiten	18
5.4.1	Team „Trusted-Health-Apps“	18
5.4.2	Wissenschaftliche Redaktion	18
5.4.2	Gutachten-Board Medizin (GBM)	18
5.4.4	Indikationsspezifische Arbeitsgruppe (iAG)	19
5.4.5	Gutachterinnen und Gutachter	19
5.5	Gutachten	20
5.6	Ermittlung der Ergebnisse	21
5.7	Aufbereitung und Kommunikation der Ergebnisse	23
5.8	Aktualisierungen	23
6	Rahmenbedingungen	25
7	Ausblick	26
	Impressum	27

Abkürzungen

DiGA	Digitale Gesundheitsanwendungen
DVG	Digitale-Versorgung-Gesetz
DVPMG	Digitale-Versorgung-und-Pflege-Modernisierungs-Gesetz
DiPA	Digitale Pflegeanwendungen
GBM	Gutachten-Board Medizin
MPG	Medizinproduktegesetz
PICO-Schema	P = Population oder Patient, Problem, I = Intervention, C = Comparison oder Control (Vergleichsintervention), O = Outcome / Endpunkt
pSVV	patientenrelevante Struktur- und Verfahrensverbesserungen

Tabellen

TABELLE 1: Schema der strukturierten Übersicht „Medizinisch-fachliche Fundierung“	16
TABELLE 2: Schema der strukturierten Übersicht „Qualität der Evidenz der positiven Versorgungseffekte“ von Primärstudien	16
TABELLE 3: Schema der strukturierten Übersicht „Qualität der Evidenz der positiven Versorgungseffekte“ von systematischen Reviews	17
TABELLE 4: Schema der Evidenzgraduierung	17
TABELLE 5: Bewertungsschema „Medizinisch-fachliche Fundierung“	22
TABELLE 6: Bewertungsschema „Qualität der Evidenz der positiven Versorgungseffekte“	22

Abbildungen

ABBILDUNG 1: Struktur und Prozess der Begutachtung	15
ABBILDUNG 2: Prozessschritte und Zuständigkeiten	18
ABBILDUNG 3: Exemplarische Ergebniskarte	24

1 Einleitung und Zusammenfassung

Das vorliegende Dokument beschreibt die Methode zur Begutachtung digitaler Gesundheitsanwendungen (DiGA) aus medizinischer Perspektive im Projekt „Trusted-Health-Apps“. Erfasst werden sowohl DiGA im Sinne von §§ 33a und 139e SGB V als auch weitere Gesundheits-Apps mit Zulassung als Medizinprodukt nach § 3 Medizinproduktegesetz (MPG). Die Methode wurde von der Bertelsmann Stiftung und ihrer Tochtergesellschaft Weisse Liste gGmbH (Weisse Liste) gemeinsam mit dem dafür eingerichteten Gutachten-Board Medizin (GBM) entwickelt. Das Dokument wird mit Weiterentwicklung der Methode kontinuierlich angepasst.

Folgende Aspekte sind für das Projekt und die Methode von besonderer Bedeutung:

Zielsetzung Qualitätstransparenz

Ziel des Projekts „Trusted-Health-Apps“ ist, die Qualität digitaler Gesundheitsanwendungen transparent zu machen. Mit diesem Hebel will das Projekt mittelfristig dazu beitragen, nutzenstiftende Anwendungen in den Versorgungsalltag zu integrieren und so die Versorgung zu verbessern. Die Ergebnisse der Begutachtung sollen Ärztinnen und Ärzten, die mit der Verordnung von DiGA beauftragt sind, als Orientierungshilfe dienen (Kapitel 2)

Nebenziel Systementwicklung

Die DiGA-Bewertung aus medizinischer Perspektive ist ein im Gesundheitssystem vergleichsweise junger Bereich. Bisher haben sich noch keine allgemein anerkannten und nachhaltigen Standards etabliert. Die Projektinitiatoren und das GBM wollen vor dem Hintergrund der praktischen Erfahrungen, die im Projekt gewonnen werden, Impulse zur (Weiter-)Entwicklung solcher Standards geben. Zudem sollen verschiedene Stakeholder konkrete Hinweise für die Qualitätsentwicklung von DiGA sowie zur Generierung von Evidenz hinsichtlich der jeweils intendierten positiven Versorgungseffekte erhalten (Kapitel 4 und 5.8).

Mehrstufiger Prozess

Die Methode besteht aus einem mehrstufigen, strukturierten Begutachtungsprozess, dessen Ziel eine unabhängige, möglichst objektive und aussagekräftige DiGA-Bewertung ist. Der Prozess soll – vor dem Hintergrund der Dynamik, welche die Entwicklung im Feld kennzeichnet – zugleich anpassungsoffen und handhabbar sein. In die Begutachtung werden externe Fachexpertinnen und Fachexperten in unterschiedlichen Rollen einbezogen. Perspektivisch wird eine enge Zusammenarbeit mit medizinischen Fachgesellschaften angestrebt (Kapitel 5).

Fokus Medizin

Übergeordnet bewertet werden zwei Aspekte: erstens die medizinisch-fachliche Fundierung von DiGA – also die Orientierung am aktuellen Stand des medizinischen Wissens – und zweitens die Qualität der Evidenz der positiven Versorgungseffekte. Bewertet wird daten- und dokumentenbasiert, nicht das Produkt selbst. Andere Qualitätsaspekte – etwa die Themen Datensicherheit oder Nutzerfreundlichkeit – werden nicht in die Bewertung einbezogen (Kapitel 5.3, 5.4, 5.6 und 5.7).

Methodentransparenz

Die Methode wird Schritt für Schritt weiterentwickelt. Der Anspruch ist eine nachvollziehbare und für alle Seiten faire DiGA-Bewertung. Entsprechend werden sowohl der jeweilige Entwicklungsstand der Methode als auch mögliche Interessenkonflikte offengelegt. Die Weiterentwicklung erfolgt im kontinuierlichen Dialog mit der Fachöffentlichkeit und den Anspruchsgruppen (Kapitel 6).

2 Projekthintergrund

Unter dem Titel „Trusted-Health-Apps“ schaffen die Bertelsmann Stiftung und die Weisse Liste Datengrundlagen und Informationsangebote, um die Qualität von DiGA transparent zu machen.

Mit dem Hebel Transparenz möchte das Projekt mittelfristig dazu beitragen, nutzenstiftende digitale Gesundheitsanwendungen in den Versorgungsalltag zu integrieren und dadurch die Versorgung zu verbessern. Digitale Technologien sollen zu einer breiten gesellschaftlichen Teilhabe an Gesundheit führen.

Das Projekt „Trusted-Health-Apps“ bestand ursprünglich aus verschiedenen Bausteinen zur Qualitätserhebung und -bewertung bei DiGA. Seit Ende 2021 liegt der Fokus jedoch allein auf der medizinischen Begutachtung bzw. der Etablierung einer entsprechenden funktionalen Methode. Dies wird als ein zentraler Ansatzpunkt gesehen, um DiGA in die Versorgung zu integrieren.

Im Kern geht es darum, Ärztinnen und Ärzten bei der Verordnung von DiGA eine Orientierungshilfe in Form der Bewertungsergebnisse an die Hand zu geben – und so indirekt dafür zu sorgen, dass die Anwendungen von Patientinnen und Patienten genutzt werden. Hintergrund ist die – durch verschiedene Analysen¹ gedeckte – Annahme, dass Versicherte DiGA vor allem dann nutzen, wenn diese von Gesundheitsfachkräften empfohlen werden.

Die Bertelsmann Stiftung und die Weisse Liste werden in der Entwicklung der Methode unterstützt und beraten vom GBM, das die Methode initial (mit-)entwickelt hat und im weiteren Prozess als eine Art Kontrollgremium für deren Absicherung und Weiterentwicklung sorgt.

Das GBM hat im April 2020 seine Arbeit aufgenommen. Es besteht aus ärztlichen und psychologischen Expertinnen und Experten verschiedener Fachrichtungen. Die Mitglieder erstellen selbst keine Gutachten, sondern begleiten den Begutachtungsprozess.

Dem GBM gehören aktuell folgende Personen an:

- Jun.-Prof. Dr. Maike Henningsen, Juniorprofessur für Digitale Medizin, Universität Witten/Herdecke, Fachärztin für Gynäkologie und Geburtshilfe
- Jun.-Prof. Dr. Stefanie Jungmann, Juniorprofessur für Klinische Psychologie & Psychotherapie des Kindes- und Jugendalters, Universität Mainz
- Prof. Dr. Jochen Klucken, Professur für Digital Medicine, Universität Luxemburg, Facharzt für Neurologie
- Prof. Dr. Sebastian Kuhn, Professur für Digitale Medizin, Universität Bielefeld, Facharzt für Orthopädie & Unfallchirurgie (Vorsitzender GBM)
- Sabine Maur, Praxis für Verhaltenstherapie, Mainz, Psychologische Psychotherapeutin

1 Mayer, M., Nöthen, T. (2021): Körper-Bürgerdelphi Gesundheitsdaten. Wie gehen wir mit elektronischer Patientenakte, Gesundheits-Apps und Datenspenden für die Forschung um? Hamburg, Körper Stiftung, S. 8; Sachverständigenrat zur Begutachtung der Entwicklung im Gesundheitswesen (2021): Ziele und Rahmenbedingungen eines dynamisch lernenden Gesundheitssystems. Gutachten 2021, S. 186–189.

-
- Prof. Dr. Martin Möckel, Ärztlicher Leiter Notfallmedizin/zentrale Notaufnahmen und Chest Pain Units Campus Charité Mitte und Campus Virchow-Klinikum, Professor für kardiovaskuläre Prozessforschung, Charité – Universitätsmedizin Berlin, Facharzt für Innere Medizin, Nephrologie, Kardiologie und Internistische Intensivmedizin; ZWB Notfallmedizin und Klinische Notfall- und Akutmedizin
 - Dr. Olaf Schmidt, Pneumologische Gemeinschaftspraxis, Koblenz, Facharzt für Innere Medizin
 - Prof. Dr. Christoph Schöbel, Professur für Schlaf- und Telemedizin, Universität Duisburg-Essen, Facharzt für Innere Medizin
 - Dr. Robin Weidemann, Geschäftsleitung Universitätsklinikum Carl Gustav Carus, Dresden, Facharzt für Innere Medizin

Die hier beschriebene Methode ist 2021 in einer Pilotphase erfolgreich getestet und diesen Erfahrungen entsprechend adaptiert worden. Ab Anfang 2022 soll sie im Dauerbetrieb erprobt werden.

Für die operative Projektsteuerung sind folgende Personen zuständig:

- Sina Busse, Weisse Liste gGmbH
- Dr. Stefan Etgeton, Bertelsmann Stiftung
- Stefanie Hennig, Bertelsmann Stiftung
- Dr. Johanna Ludwig, Unfallkrankenhaus Berlin, Fachärztin für Orthopädie & Unfallchirurgie (Medizinische Beraterin Bertelsmann Stiftung)
- Timo Thranberend, Bertelsmann Stiftung (bis Ende Oktober 2021)

3 Definitionen

Im Folgenden werden zunächst Begriffe erläutert, die im Zusammenhang mit der DiGA-Begutachtung aus medizinischer Perspektive im Projekt „Trusted-Health-Apps“ relevant sind.

3.1 Digitale Gesundheitsanwendungen (DiGA)

DiGA sind der zentrale Gegenstand der Begutachtung.

Bereits seit einigen Jahren werden Gesundheits-Apps, die als Medizinprodukt nach § 3 MPG zugelassen sind, in begrenztem Umfang in die medizinischen und psychotherapeutischen Behandlungsabläufe integriert. Dabei unterscheiden sich die Rahmenbedingungen bzw. die Vertrags- und Vergütungsvarianten, unter denen die Anwendungen eingesetzt und erstattet werden. Relativ häufig zu beobachten ist der Einsatz im Rahmen von Selektivverträgen nach § 140 SGB V bzw. im Kontext von Prävention und Gesundheitsförderung nach § 20 SGB V.

Mit dem Inkrafttreten des Digitale-Versorgung-Gesetzes (DVG) im Dezember 2019 ist eine neue, für das Feld besonders wichtige Vertrags- und Erstattungsoption hinzugekommen. Seitdem können Ärztinnen und Ärzte bzw. Psychotherapeutinnen und Psychotherapeuten DiGA verordnen, und gesetzlich Versicherte haben den Anspruch, sich die Kosten von ihrer Krankenkasse erstatten zu lassen. Eine DiGA wird dann ordnungsfähig bzw. zur „App auf Rezept“,

wenn sie definierte Anforderungen erfüllt sowie ein Prüfverfahren beim Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) durchlaufen hat. Nach erfolgreicher Prüfung wird sie im sogenannten DiGA-Verzeichnis (<https://diga.bfarm.de>) geführt. Dies erfolgte erstmals im Oktober 2020.

DiGA im Sinne des Gesetzes (§§ 33a und 139e SGB V) bzw. des DiGA-Verzeichnisses vom BfArM sind derzeit definiert als Medizinprodukte, die folgende Eigenschaften aufweisen:²

- Medizinprodukt der Risikoklasse I oder IIa nach Medical Device Regulation (MDR) (oder im Rahmen der Übergangsvorschriften nach Medical Device Directive (MDD))
- Ihre Hauptfunktion beruht auf digitalen Technologien (App oder browserbasierte Anwendung).
- Der medizinische Zweck wird wesentlich durch die digitale Hauptfunktion erreicht.
- Sie unterstützen die Erkennung, Überwachung, Behandlung oder Linderung von Krankheiten oder die Erkennung, Behandlung, Linderung oder Kompensierung von Verletzungen oder Behinderungen (keine Primärprävention).
- Sie werden von Patientinnen und Patienten oder von Leistungserbringer und Patient gemeinsam genutzt.

Gemäß dem „Digitale-Versorgung-und-Pflege-Modernisierungs-Gesetz“ (DVPMG) gibt es künftig zudem für sogenannte „Digitale Pflegeanwendungen“ (DiPA) ein ähnliches Prüfver-

² Bundesinstitut für Arzneimittel und Medizinprodukte (Hrsg.): Das Fast Track Verfahren für digitale Gesundheitsanwendungen nach § 139e SGB V. Leitfaden für Hersteller, Leistungserbringer und Anwender (Stand: 23.10.2020).

fahren wie bei DiGA und einen Erstattungsweg über die gesetzliche Pflegeversicherung. Ebenso wurde 2020 die Prüfung von Präventionsanwendungen für eine Erstattung in der gesetzlichen Krankenversicherung vom GKV-Spitzenverband im Leitfaden Prävention³ neu geregelt.

Das Projekt „Trusted-Health-Apps“ schließt alle digitalen Anwendungen ein, unabhängig von den Rahmenbedingungen, unter denen sie eingesetzt werden. Primäres Abgrenzungsmerkmal ist die Zertifizierung als Medizinprodukt. Entsprechend gelten folgende Einschlusskriterien:

- Digitale Anwendung (z. B. native Smartphone-, Tablet- oder Smartwatch-Anwendung, Webanwendung oder Sprachanwendung für Sprachassistenten)
- Zertifiziertes Medizinprodukt – nach MDR oder, im Rahmen der Übergangsvorschriften, nach MDD (unabhängig von der Risikoklasse)
- Benutzeroberfläche für Patientinnen und Patienten sowie Versicherte (Nutzung durch Patientinnen und Patienten oder gemeinsam durch Patientinnen und Patienten sowie Angehörige von Gesundheitsberufen)
- Hauptfunktion beruhend auf digitalen Softwaretechnologien der digitalen Gesundheitsanwendung selbst (Erreichung des medizinischen Zwecks durch digitale Hauptfunktion, nicht nur durch Auslesen oder Steuerung eines anderen Medizingeräts)
- Deutschsprachig (oder Vorhandensein einer deutschsprachigen Version)

In Abgrenzung hierzu existieren etliche Gesundheits-Apps ohne Zertifizierung als Medizinprodukt. Diese werden im Projekt „Trusted-Health-Apps“ nicht betrachtet.

Hinweis: Der Begriff „Digitale Gesundheitsanwendung“ (DiGA) wird derzeit nicht einheitlich verwendet. In der Fach-Community setzt sich aktuell eine Verwendung durch, die aus dem Rechtsbegriff im SGB V resultiert: DiGA sind demnach nur solche Anwendungen, die das BfArM erfolgreich geprüft hat und die im DiGA-Verzeichnis gelistet werden. Im Projekt „Trusted-Health-Apps“ gilt „DiGA“ indes als Überbegriff für alle Anwendungen, die den oben beschriebenen Einschlusskriterien des Projekts entsprechen. Diese begriffliche Unschärfe wird bewusst akzeptiert, weil derzeit kein anderer Ausdruck existiert, der eine hinreichend klare und abgrenzende Bezeichnung erlaubt.

3.2 AppQ

AppQ⁴ ist das zugrunde liegende Gütekriterien-Modell für die Datenerhebung zur Begutachtung von DiGA.

Gefördert vom Bundesministerium für Gesundheit und unter Beteiligung des Fraunhofer-Instituts für Offene Kommunikationssysteme (FOKUS), hat die Bertelsmann Stiftung 2019 ein Kernset von Gütekriterien für DiGA ent-

3 GKV-Spitzenverband (2020): Leitfaden Prävention. Handlungsfelder und Kriterien nach § 20 Abs. 2 SGB V.

4 Thranberend, T., Bittner, J. (2019): AppQ. Gütekriterien-Kernset für mehr Qualitätstransparenz bei digitalen Gesundheitsanwendungen, Gütersloh, Bertelsmann Stiftung; Thranberend, T., Bittner, J. (2020): AppQ 1.1. Gütekriterien-Kernset für mehr Qualitätstransparenz bei digitalen Gesundheitsanwendungen, Gütersloh, Bertelsmann Stiftung, S. 18.

wickelt und für Hersteller-Selbstangaben nutzbar gemacht. Das Kernset ermöglicht, die Qualität der als Medizinprodukt zertifizierten DiGA einzuschätzen und sie standardisiert zu vergleichen, unabhängig von ihrem Typ oder ihrer Risikoklasse.

AppQ besteht aus zwei wesentlichen Teilen: einem Metadatenmodell zur Beschreibung und Klassifizierung einer DiGA sowie dem eigentlichen Gütekriterien-Kernset, auf dessen Basis strukturierte Qualitätstransparenz ermöglicht wird. Das Kernset umfasst aktuell neun Themenbereiche mit insgesamt 25 Kriterien und 187 Indikatoren.

Die AppQ-Kriterien entsprechen in ihrer Ausrichtung weitgehend denen, die das BfArM in der sogenannten DiGA-Verordnung für die Datenerhebung und Prüfung vorgibt.

3.3 Medizinisch-fachliche Fundierung

Die medizinisch-fachliche Fundierung einer DiGA ist einer von zwei Bereichen, die in der Begutachtung bewertet werden.

Das GBM betrachtet im Schwerpunkt die medizinisch-fachliche Fundierung, d. h. die Leitlinien und Publikationen, die vom Hersteller bei der Entwicklung seiner DiGA zugrunde gelegt wurden. Hierbei wird geprüft, ob es einen konkreten Bezug gibt zwischen Zielsetzungen und Wirkweisen, der Quelle (Sekundärliteratur, Leitlinien) und der entsprechenden (Teil-)Funktion der DiGA. Im Mittelpunkt der Begutachtung in diesem Bereich steht somit, ob relevante Erkennt-

nisse aus wissenschaftlichen (Sekundär-)Studien und Leitlinien berücksichtigt wurden.

3.4 Qualität der Evidenz der positiven Versorgungseffekte

Der Nachweis positiver Versorgungseffekte im Sinne der Evidenz zu einer DiGA ist der zweite der beiden Bereiche, die in der Begutachtung bewertet werden.

Für die Beschreibung der Effekte von DiGA gibt es derzeit noch kein allgemein anerkanntes Begriffskonzept. Die Begriffe „Nutzen“, „Wirksamkeit“ und „Versorgungseffekt“ werden im Fachdiskurs nicht einheitlich verwendet. Als positive Versorgungseffekte im Sinne dieses Dokuments werden zum einen solche Effekte bezeichnet, welche die medizinische Wirksamkeit oder den medizinischen Nutzen von digitalen DiGA betreffen. Zum anderen umfasst der Begriff patientenrelevante Struktur- und Verfahrensverbesserungen (pSVV) in der Gesundheitsversorgung. Die in beiden Kriterien erfragten Nachweise beziehen sich ausschließlich auf Effekte, die für die jeweilige DiGA speziell erbracht wurden oder werden sollen. Neben den Nachweisen positiver Effekte durch Einsatz der DiGA wird auch nach dokumentierten Hinweisen auf unerwünschte Wirkungen gefragt.

4 Herausforderung: Medizinische Bewertung von DiGA

Für die Gesellschaft und für Gesundheitsfachkräfte stellt sich – unabhängig von der jeweiligen Vertrags- oder Vergütungsvariante einer digitalen Gesundheitsanwendung – die Herausforderung, einen Überblick über das Marktangebot zu erhalten und die Qualität der jeweiligen Anwendung einzuschätzen, vor allem im Hinblick auf den patientenrelevanten Mehrwert. Dabei werden die medizinisch-fachliche Fundierung und die positiven Versorgungseffekte einer Anwendung von den meisten für die wichtigsten Parameter gehalten.⁵

In diesem für das Gesundheitssystem vergleichsweise jungen Bereich haben sich jedoch noch keine nachhaltigen und allgemein anerkannten Standards etabliert, um die medizinisch-fachliche Fundierung und die Qualität der Evidenz positiver Versorgungseffekte nachzuweisen und den jeweiligen Nachweis extern zu überprüfen. Der geschilderte Informationsbedarf ist aktuell weitgehend ungedeckt.

Hinsichtlich der medizinisch-fachlichen Fundierung gilt bislang als ausreichend, Referenzen zu einschlägigen Leitlinien und Publikationen anzugeben. Der konkrete Zusammenhang zwischen Zielsetzung und Wirkweise auf der einen Seite und der DiGA-Umsetzung auf der anderen Seite lässt sich jedoch oft schwer erfassen.

In Bezug auf den Nachweis der Evidenz positiver Versorgungseffekte fordert das Fast-Track-Verfahren für DiGA nach § 139e SGB V vom Hersteller quantitative vergleichende Studien, deren „gewählte Methodik adäquat zum gewählten Untersuchungsgegenstand ist“. Während verschiedene wissenschaftliche Arbeitsgruppen⁶ hierzu eine systematische Festlegung anhand von Risikoprofilen empfehlen, steht die regulatorische Spezifizierung derzeit aus. Als Publikationsort der Studien gilt zunächst auch die Website der Hersteller als ausreichend; später muss der Eintrag in das Verzeichnis erfolgen. Bei digitalen Medizinprodukten nach § 3 MPG ist dieser Bereich weitestgehend unregelt.

Der Einsatz von DiGA in der medizinischen und psychotherapeutischen Versorgung ist zudem je nach Funktionsumfang als komplexe Intervention zu betrachten. Eine Qualitätsbewertung und der Wirksamkeitsnachweis digitaler Anwendungen stellen demnach sowohl die bisherigen Versorgungsstrukturen als auch die medizinische Wissenschaft vor neue Herausforderungen. Nicht zuletzt erscheinen aufgrund der starken Entwicklungsdynamik, der Komplexität und Heterogenität von DiGA bekannte Bewertungsmechanismen nicht mehr hinlänglich zu sein. Es bedarf neuer Ansätze und Methoden.

5 Thranberend, T., Bittner, J. (2019): AppQ. Gütekriterien-Kernset für mehr Qualitätstransparenz bei digitalen Gesundheitsanwendungen, Gütersloh, Bertelsmann Stiftung, S. 44 ff.

6 NICE (2019): Evidence standards framework for digital health technologies, <https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies> (abgerufen am 9.12.2021); Busse, R., et al. (2019): Anforderungen an Bewertungsverfahren und Evaluationsdesigns III, www.mig.tu-berlin.de/fileadmin/a38331600/2019.teaching.ws/I.DiGA_Ergebnispapier_3.pdf (abgerufen am 20.10.2021).

5 Der Begutachtungsprozess

5.1 Zielsetzung

Primäres Ziel des Begutachtungsprozesses ist – im Sinne des Projekts „Trusted-Health-Apps“ –, die Qualität von DiGA transparent zu machen. Dabei sollen die medizinisch-fachliche Fundierung und die Qualität der publizierten Evidenz der positiven Versorgungseffekte nachvollziehbar eingeordnet werden und somit eine Orientierung bieten. Darüber hinaus sollen mit der Begutachtung fortlaufend praktische Erkenntnisse zur oben genannten Herausforderung gewonnen und Ableitungen für regulatorische oder fachliche Weiterentwicklungen getroffen werden, etwa mit Blick auf gesetzliche Vorgaben oder das Design von Studien zu DiGA. Um der geschilderten Komplexität der DiGA-Bewertung gerecht zu werden, wird ein mehrstufiger, strukturierter und zugleich handhabbarer wie anpassungsoffener Begutachtungsprozess implementiert, der im Folgenden dargelegt wird.

5.2 Bewertungsgegenstand

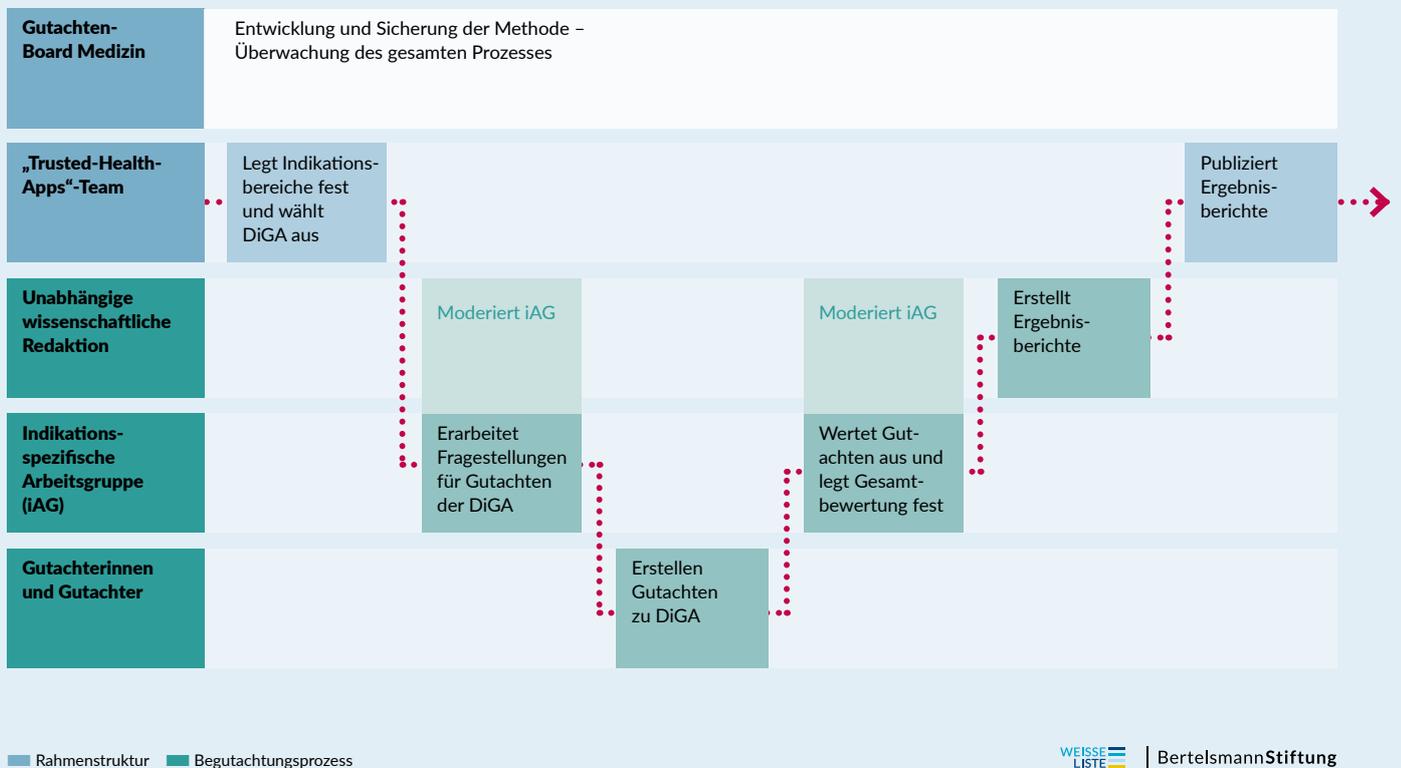
Gegenstand der Begutachtung bzw. Bewertung sind DiGA, also Anwendungen, die den oben beschriebenen Kriterien entsprechen. Durch die Begutachtung sollen insbesondere Aussagen über die Qualität der medizinischen Evidenz gewonnen werden, die den zu untersuchenden DiGA zugrunde liegt. Nutzungsbezogene Eigenschaften der DiGA, etwa deren Nutzerfreundlichkeit, werden jedoch nicht berücksichtigt. Die Bewertung erfolgt ausschließlich anhand der Daten-, Dokumenten- und Studienlage und nicht anhand einer praktischen Testung.

Die DiGA werden üblicherweise innerhalb von Projekten bewertet, die mehrere Anwendungen bzw. spezifische Module oder Komponenten von Anwendungen eines definierten Indikationsbereichs (z. B. „Depressionen“) bzw. einer Produktgruppe (z. B. „Symptom-Checker“) umfassen. Welche Bewertungsprojekte durchgeführt werden, entscheidet das Team des Projekts „Trusted-Health-Apps“ in Abstimmung mit dem GBM. Die Festlegung erfolgt anhand qualitativer Maßstäbe wie Markt- bzw. Versorgungsrelevanz bzw. aufgrund eines besonderen (z. B. wissenschaftlichen) Erkenntnisinteresses, das an die Bertelsmann Stiftung, die Weisse Liste oder das GBM herangetragen wurde.

Innerhalb eines Bewertungsprojekts werden wiederum nicht sämtliche Anwendungen eines Indikationsbereichs bzw. einer Produktgruppe betrachtet, sondern lediglich ausgewählte. Hintergrund dafür sind Erfahrungen aus der Pilotphase: Hier wurde deutlich, dass der Begutachtungsprozess sowohl in der Gesamtsteuerung durch das Team „Trusted-Health-Apps“ als auch für die Gutachterinnen und Gutachter nur mit einer begrenzten Zahl von Anwendungen pro Bereich oder Gruppe sinnvoll realisierbar ist.

Die Auswahl der Anwendungen für ein Projekt erfolgt ebenfalls qualitativ. Dabei werden entweder besonders homogene Gruppen gebildet (z. B. anhand der Funktionsweise, der intendierten Versorgungseffekte oder des eingesetzten Studiendesigns) oder – im Gegenteil – bewusst unterschiedliche Anwendungen betrachtet. Ein Anspruch von DiGA-Herstellern auf Begutachtung besteht nicht. Allerdings ist es Ziel, sowohl das Spektrum der begutachteten DiGA als auch die Anzahl mit der Durchführung mehrerer Bewertungsprojekte pro Jahr sukzessive zu steigern.

ABBILDUNG 1: Struktur und Prozess der Begutachtung



5.3 Ausgangsbasis

Die Begutachtung von DiGA erfolgt übergeordnet anhand von zwei Leitfragen.

Welche medizinisch-fachliche Fundierung hat die begutachtete DiGA und inwieweit beruht sie auf dem aktuellen Stand medizinischen Wissens im Fachgebiet? (Medizinisch-fachliche Fundierung, Tabelle 1)

Sind die intendierten positiven Versorgungseffekte der begutachteten DiGA nachgewiesen und wie gut ist die Qualität dieser Nachweise? (Qualität der Evidenz der positiven Versorgungseffekte, Tabelle 2 & Tabelle 3)

Als Ausgangsbasis der Begutachtung bereitet das Team „Trusted-Health-Apps“ für jeden der zwei Bereiche eine strukturierte Übersichtstabelle vor – zum einen mit Daten, die bereits in der Daten-

bank der Redaktion vorhanden sind (Selbstrecherchen sowie Daten aus der DiGA-API⁷ des BfArM), zum anderen über gezielte Erhebungen bei den Herstellern für die jeweilige Begutachtung. Alle Hersteller erhalten dabei die Möglichkeit, bis zu einem definierten Stichtag innerhalb eines Bewertungsprojekts zu folgenden Fragen Stellung zu nehmen:

- Bitte beschreiben Sie in eigenen Worten, was [Name der DiGA] ist und wie die Anwendung eingesetzt wird.
- An welche Zielgruppe richtet sich Ihre Anwendung konkret und für welchen Anwendungsbereich (ICD-10 Codes) ist sie vorgesehen?
- Was ist die Hauptfunktion Ihrer Anwendung? Gibt es weitere Teilfunktionen?
- Wie lange ist eine Nutzung der Anwendung üblicherweise vorgesehen?
- Welche medizinischen Zielsetzungen und

7 <https://simplifier.net/guide/diga/home>

Wirkweisen verfolgt Ihre Anwendung, aus welchen Quellen (Primär- / Sekundärliteratur, Leitlinien) sind diese abgeleitet und mit welchen (Teil-)Funktionen Ihrer Anwendung werden die Zielsetzungen verfolgt? Um eine möglichst aussagekräftige Beurteilung erstellen zu können, benötigen die Gutachterinnen und Gutachter konkrete Referenzen (Seitenzahl, Kapitel) in den bereitgestellten Quellen.

- Gibt es dokumentierte Hinweise auf unerwünschte medizinische Wirkungen der Anwendung auf Nutzerinnen und Nutzer?
- Gibt es Unterstützungsmechanismen, die Nutzerinnen und Nutzern bei medizinischen Problemen zur Verfügung stehen?

- Was sind die intendierten positiven Versorgungseffekte Ihrer Anwendung?
- Folgende Primärstudien liegen uns für die Beurteilung der Qualität der Evidenz der positiven Versorgungseffekte durch die Gutachterinnen und Gutachter vor. Ist diese Aufzählung vollständig? [Listung der Studien]
- Sind eine oder mehrere Studien, z. B. aufgrund fehlender Aktualität der Erhebung, nicht in die Bewertung einzubeziehen?

TABELLE 1: Schema der strukturierten Übersicht „Medizinisch-fachliche Fundierung“

DiGA	Name der DiGA oder digitalen Lösung inkl. Website-URL
Hersteller-Kurzbeschreibung	Kurzbeschreibung der DiGA des Herstellers
Anwendungsbereich nach ICD-10	Codes
Hauptfunktion	Hauptfunktion der DiGA
Zielgruppe und vorgesehene Dauer der Nutzung	Zielgruppe nach Alter, Geschlecht, Gesundheitszustand und Versorgungsstufe sowie vorgesehene Nutzungsdauer
Zielsetzung, Wirkweise, Inhalt und Nutzung	Benennung und Beschreibung
Quellen	Primär- / Sekundärliteratur und Leitlinien mit konkreter Referenz (Seitenzahl, Kapitel) und Link
Funktion der DiGA	Teilfunktion der DiGA zur Erreichung der Zielsetzung

TABELLE 2: Schema der strukturierten Übersicht „Qualität der Evidenz der positiven Versorgungseffekte“ von Primärstudien

DiGA	Name der DiGA oder digitalen Lösung inkl. Website-URL
Hersteller-Kurzbeschreibung	Kurzbeschreibung der DiGA des Herstellers
Anwendungsbereich nach ICD-10	Codes
Hauptfunktion	Hauptfunktion der DiGA
Studie / Publikation	Zitation und Link
Studientyp	Studiendesigns aus Liste (AppQ-VE-01)
Patient / Patientin	Anzahl, Diagnose, Altersgruppe, Drop-out Rate
Intervention	Art der Intervention
Comparison	Art der Vergleichsgruppe / Vergleichsgruppen-Intervention
Outcome 1 / Primärer Endpunkt (medizinischer Nutzen)	Intendierter Versorgungseffekt, konkreter Endpunkt, Surrogatparameter und Methode
Outcome 2 / Weitere Endpunkte (sonstige positive Versorgungseffekte)	Intendierter Versorgungseffekt, konkreter Endpunkt, Surrogatparameter und Methode
Evidenzgraduierung	Evidenzlevel* (Tabelle 4)

* Scottish Intercollegiate Guidelines Network (2001); SIGN 50: a guideline developers' handbook, Edinburgh.

Die tabellarische Übersicht zur Qualität der Evidenz der positiven Versorgungseffekte beinhaltet das PICO-Schema. Erfasst werden alle Primärstudien, die bisher durchgeführt und publiziert wurden (Tabelle 2) und systematische Reviews (Tabelle 3). Das Schema der Evidenzgraduierung nach SIGN⁸ bildet die Grundlage für die Bestimmung des Evidenzlevels (Tabelle 4).

Die Übersichten werden ergänzt durch Links zu den genannten Publikationen und Leitlinien oder entsprechende Dokumente als PDF. Zudem fertigt die wissenschaftliche Redaktion eine

Gesamtübersicht zu allen DiGA im jeweiligen Bewertungsprojekt an und ordnet diese im Kontext des jeweiligen Marktsegments bzw. des Indikationsbereichs ein. Dieses „Materialpaket“ wird der jeweiligen indikationsspezifischen Arbeitsgruppe (iAG, Kapitel 5.4.4) wie auch den Gutachterinnen und Gutachtern zur Verfügung gestellt – per E-Mail oder perspektivisch über eine entsprechende interne Website des Projekts „Trusted-Health-Apps“.

TABELLE 3: Schema der strukturierten Übersicht „Qualität der Evidenz der positiven Versorgungseffekte“ von systematischen Reviews

DiGA	Name der DiGA oder digitalen Lösung inkl. Website-URL
Hersteller-Kurzbeschreibung	Kurzbeschreibung der DiGA des Herstellers
Anwendungsbereich nach ICD-10	Codes
Hauptfunktion	Hauptfunktion der DiGA
Studie / Publikation	Zitation und Link
Studientyp	Studiendesigns aus Liste (AppQ-VE-01)
Patienten- / Studienzahl	Patientenzahl (gesamt) Studienzahl
Intervention	Art der Intervention(en)
Endpunkte	Erhobene Endpunkte
Ergebnis	Ergebniszusammenfassung
Evidenzgraduierung	Evidenzlevel (Tabelle 4)

TABELLE 4: Schema der Evidenzgraduierung (nach SIGN)

Evidenzlevel	Einordnung
1++	Qualitativ hochwertige Metaanalysen, Systematische Übersichten von RCTs, oder RCTs mit sehr geringem Risiko systematischer Fehler (Bias)
1+	Gut durchgeführte Metaanalysen, Systematische Übersichten von RCTs, oder RCTs mit geringem Risiko systematischer Fehler (Bias)
1-	Metaanalysen, Systematische Übersichten von RCTs, oder RCTs mit hohem Risiko systematischer Fehler (Bias)
2++	Qualitativ hochwertige systematische Übersichten von Fall-Kontroll- oder Kohortenstudien oder qualitativ hochwertige Fall-Kontroll- oder Kohortenstudien mit sehr niedrigem Risiko systematischer Verzerrungen (Confounding, Bias, „Chance“) und hoher Wahrscheinlichkeit, dass die Beziehung ursächlich ist
2+	Gut durchgeführte Fall-Kontroll-Studien oder Kohortenstudien mit niedrigem Risiko systematischer Verzerrungen (Confounding, Bias, „Chance“) und moderater Wahrscheinlichkeit, dass die Beziehung ursächlich ist
2-	Fall-Kontroll-Studien oder Kohortenstudien mit einem hohen Risiko systematischer Verzerrungen (Confounding, Bias, „Chance“) und signifikantem Risiko, dass die Beziehung nicht ursächlich ist
3	Nicht-analytische Studien, z. B. Fallberichte, Fallserien
4	Expertenmeinung

8 Scottish Intercollegiate Guidelines Network (2001): SIGN 50: a guideline developers’ handbook, Edinburgh.

ABBILDUNG 2: Prozessschritte und Zuständigkeiten

Prozessschritt	Team „Trusted-Health-Apps“	Wissenschaftliche Redaktion	GBM gesamt	GBM-Vorsitz
1 Festlegung Bewertungsprojekt				Bestätigung
2 DiGA-Auswahl				Bestätigung
3 Datenerhebung & Aufbereitung	Unterstützung			
4 Konstitution iAG			Vorschläge	Bestätigung
5 Durchführung iAG		Moderation	Teilnahme Einzelner	
6 Auswahl und Akquise Gutachter				Bestätigung
7 Beauftragung & Steuerung Gutachten				
8 Durchführung iAG		Moderation	Teilnahme Einzelner	
9 Erstellung Ergebnisformate				
10 Dialog mit Herstellern (opt.)		Fachliche Unterstützung		
11 Veröffentlichung Ergebnisse			Kommentierung	
12 Überprüfung Methode				
13 Stakeholder-Dialog				

■ volle Zuständigkeit ■ arbeitsteiliger Vorgang

WEISSE LISTE | BertelsmannStiftung

5.4 Rollen und Verantwortlichkeiten

Der Begutachtungsprozess erfolgt im Zusammenspiel verschiedener Akteurinnen und Akteure. Im Folgenden werden die jeweiligen Rollen und Verantwortlichkeiten beschrieben.

5.4.1 Team „Trusted-Health-Apps“

Das Team „Trusted-Health-Apps“ der Bertelsmann Stiftung und der Weissen Liste steuert den gesamten Begutachtungsprozess operativ – von der Festlegung eines Bewertungsprojekts bis zu der Veröffentlichung der Ergebnisse und der Überprüfung der Methode sowie der Erstellung eines Whitepapers zu den übergeordneten Erkenntnissen eines Projekts (siehe Abbildung 2). Das Team übernimmt die Kommunikation mit allen Beteiligten im Prozess und stimmt sich dabei mit dem GBM (Kapitel 5.4.3) ab.

5.4.2 Wissenschaftliche Redaktion

Die wissenschaftliche Redaktion stellt die fachliche Korrektheit in den operativen Prozessschritten wie der Datenaufbereitung sowie der Aufbereitung der Ergebnisse sicher, moderiert die iAG, erstellt den Ergebnisbericht und unterstützt den Dialog mit den Herstellern über die Ergebnisse.

5.4.3 Gutachten-Board Medizin (GBM)

Das GBM besteht aus Ärztinnen und Ärzten sowie Psychotherapeutinnen und Psychotherapeuten mit ausgewiesener Expertise im Kontext digitaler Gesundheitsanwendungen. Sie haben die in diesem Dokument beschriebene Methode maßgeblich (mit-)entwickelt. Im weiteren Verlauf soll das Gremium die Durchführung der Methode absichern und diese zusammen mit dem Team „Trusted-Health-Apps“ sowie der wissenschaftlichen Redaktion regelmäßig evaluieren und, wenn nötig, iterativ weiterent-

wickeln. Zudem übernimmt das GBM eine besondere Rolle im Stakeholder-Dialog zu den in den Bewertungsprojekten gewonnenen übergeordneten Erkenntnissen.

In einem Bewertungsprojekt arbeitet jeweils mindestens ein Mitglied in der iAG mit. Zudem erhalten alle Mitglieder die Möglichkeit, die Ergebnisse vor der Veröffentlichung zu kommentieren und damit auf potenziell notwendige fachliche oder methodische Korrekturen hinzuweisen.

Aus der Mitte des Boards wird – nach Vorschlag der Bertelsmann Stiftung und der Weissen Liste – jeweils ein Board-Vorsitz bestimmt. Diesen Vorsitz haben eine oder zwei Personen, die dem Team „Trusted-Health-Apps“ als beratende Instanz zur Verfügung stehen sowie als Ansprechpersonen für die Bestätigung von Bewertungsprojekt, DiGA-Auswahl, Zusammensetzung der iAG sowie Auswahl der Gutachterinnen und Gutachter. Der Vorsitz wird üblicherweise mindestens für ein Jahr bestimmt.

5.4.4 Indikationsspezifische Arbeitsgruppe (iAG)

Eine iAG besteht aus mindestens einem GBM-Mitglied und in der Regel zwei weiteren Personen. Hinzukommen sollten digital-affine Ärztinnen und Ärzte bzw. Psychotherapeutinnen und Psychotherapeuten mit klinischer und/oder wissenschaftlicher Expertise aus dem jeweiligen Indikationsgebiet des Bewertungsprojekts. Auch weitere Gesundheitsfachberufe, etwa aus den Feldern Physiotherapie oder Pflege, können in der Arbeitsgruppe vertreten sein. Zudem ist ein Mitglied des Teams „Trusted-Health-Apps“ beratend eingebunden. Moderiert werden die iAG-Sitzungen von der wissenschaftlichen Redaktion.

Die Expertinnen und Experten der einzelnen iAGs werden namentlich genannt. Perspektivisch können deren Mitglieder ihre jeweilige (medizinische) Fachgesellschaft repräsentieren. Zunächst sind sie jedoch, ebenso wie die GBM-Mitglieder, als Person, nicht als Repräsentant oder Repräsentantin einer Organisation in der Arbeitsgruppe vertreten.

In der ersten Sitzung werden die tabellarischen Übersichten zur medizinisch-fachlichen Fundierung und zur Qualität der Evidenz der positiven Versorgungseffekte gesichtet. Dabei werden zum einen mögliche Ergänzungserfordernisse und Korrekturen herausgearbeitet; zum anderen geht es um die Identifizierung und Formulierung von „Key Uncertainties“, also Unklarheiten mit Blick auf die Einordnung der vorliegenden Daten (z. B. in Bezug auf Studienendpunkte oder gewählte Beobachtungszeiträume). Diese „Key Uncertainties“ werden anschließend von der wissenschaftlichen Redaktion als Fragestellungen für die Gutachten ausgearbeitet.

In der zweiten Sitzung werden die Gutachten ausgewertet. Dies erfolgt zum einen qualitativ, zum anderen in Form definierter Skalen für die beiden Bewertungsbereiche (medizinisch-fachliche Fundierung und Qualität der Evidenz der positiven Versorgungseffekte) sowie für eine Gesamtbewertung (Kapitel 5.7). Sollten die beiden unabhängigen Gutachten größeren Dissens aufweisen, kann die iAG empfehlen, ein drittes Gutachten zu beauftragen.

5.4.5 Gutachterinnen und Gutachter

Die Gutachterinnen und Gutachter sind unabhängig und damit keine Mitglieder des GBM oder einer iAG. Sie sind ebenfalls digital-affine Ärztinnen und Ärzte oder Psychotherapeutinnen und Psychotherapeuten oder repräsentieren andere Gesundheitsfachberufe. Sie haben die Expertise für ein spezifisches Fachgebiet. Relevante Kriterien zu ihrer Benennung sind formale Qualifikationen wie etwa ein Facharztstatus, Schwerpunktkompetenzen und Zusatzweiterbildungen. Zudem können Benennungen aufgrund „fachgebietsübergreifender Expertise“ erfolgen (z. B. im Bereich von Gesundheitsförderung und Prävention, Disease-Management-Programmen oder Expertise im Bereich von Funktionstypologien).

Um eine objektive Begutachtung zu ermöglichen und Diskriminierung zu vermeiden, sollen die Gutachterinnen und Gutachter zunächst anonym bleiben.

Aufgabe der Gutachterinnen und Gutachter ist die schriftliche Begutachtung der DiGA aus einem Bewertungsprojekt anhand der zur Verfügung gestellten Daten und nach einem strukturierten Schema. Zudem nehmen sie Stellung zu den von der iAG formulierten „Key Uncertainties“.

5.5 Gutachten

Pro DiGA werden zwei unabhängige Gutachten nach einem strukturierten Schema erstellt. Vorgegebene Fragen und definierte Antwort-Umfänge (i. S. e. maximalen Zeichenzahl pro Antwort) sollen für möglichst objektive und vergleichbare Einschätzungen sorgen. Grundlage der Bewertung sind die oben beschriebenen Übersichten und zugehörigen Dokumente („Materialpaket“, Kapitel 5.3). Die Gutachterinnen und Gutachter erhalten, wenn möglich, zudem einen Testzugang zu den jeweiligen Anwendungen, um sich ein möglichst vollständiges Bild machen zu können. Die Anwendung selbst ist jedoch nicht Gegenstand der Bewertung, sondern nur die „Daten- und Dokumentenlage“.

Das jeweilige Gutachten deckt zwei Bereiche ab:

Bewertung der medizinisch-fachlichen Fundierung

- Sind die vom Hersteller angegebenen Bezüge zwischen Zielsetzung, Wirkweise und /oder Nutzen der DiGA anhand der angegebenen Leitlinien und Publikationen plausibel herleitbar?
- Sind die intendierte Wirkweise und die dafür umgesetzten (Teil-)Funktionen nach aktuellem Stand des Wissens grundsätzlich geeignet, die Zielsetzung zu erreichen?
- Wird die Patientensicherheit adäquat sichergestellt?

Bewertung der Qualität der Evidenz der positiven Versorgungseffekte

- Sind die vom Hersteller intendierten positiven Versorgungseffekte mit der vorgelegten Evidenz hinreichend belegt?
 - Population: Halten Sie die gewählten Studiengruppen für adäquat?
 - Intervention: Halten Sie die gewählten Interventionen für adäquat?
 - Comparison: Halten Sie die gewählten Kontrollinterventionen für adäquat?
 - Outcome: Halten Sie die jeweiligen primären oder sekundären Endpunkte für adäquat gewählt?
- Welche methodischen Ansätze halten Sie für geeignet, die Evidenz weiter zu erhöhen?
- Welche Relevanz haben die intendierten positiven Versorgungseffekte an sich aus fachlicher Perspektive für die Patientenversorgung im Fachgebiet?
- Hinweise zu möglichen unerwünschten Wirkungen
 - Wurden Nebenwirkungen im Rahmen der Studien (methodisch) adäquat untersucht?
 - Sehen Sie in den Daten weitere Hinweise auf mögliche unerwünschte Wirkungen?

Ergänzt werden diese standardisierten Fragen jeweils durch spezifische Fragen, abgeleitet aus den von den iAGs formulierten „Key Uncertainties“.

Die Fragen werden immer qualitativ mit Text beantwortet. Zusätzlich soll künftig getestet werden, ob eine ergänzende skalenbasierte Einschätzung hilfreich ist, um das Gesamtergebnis der iAG abzuleiten. So sollen alle entsprechenden Fragen ergänzt werden durch eine vierstufige Antwortskala („trifft voll und ganz zu“, „trifft eher zu“, „trifft eher nicht zu“, „trifft nicht zu“). Darüber hinaus werden die Gutachterinnen und Gutachter um aggregierte Einschätzungen gebeten zu den Bewertungsbereichen „Medizinisch-fachliche Fundierung“ und „Qualität der Evidenz der positiven Versorgungseffekte“ sowie bezüglich eines möglichen Gesamtergebnisses. Die Skalen bzw. Bewertungsschemata entsprechen jeweils denen, die bei der Ermittlung der Ergebnisse eingesetzt werden (Kapitel 5.6 und 5.7).

5.6 Ermittlung der Ergebnisse

Die Bewertungsergebnisse für die jeweilige DiGA werden in der zweiten iAG-Sitzung durch Auswertung der Gutachten ermittelt. Dabei werden zum einen die beiden zentralen Bereiche „Medizinisch-fachliche Fundierung“ und „Qualität der Evidenz der positiven Versorgungseffekte“ bewertet, zum anderen erfolgt eine Gesamtbewertung.

Die Bereiche werden qualitativ bewertet sowie auf Basis der in den folgenden Tabellen dargestellten Schemata. Diese sind abgeleitet aus gängigen medizin-wissenschaftlichen Bewertungsschemata.⁹ Die Bewertung anhand der Schemata wird nicht 1:1 veröffentlicht, sondern wird innerhalb der iAG konsentiert.

Die Gesamtbewertung dient dazu, das Ergebnis von der zentralen Zielgruppe der medizinischen Begutachtung – also Gesundheitsfachkräften, die DiGA empfehlen und verordnen – einschätzen zu lassen. Sie wird, angelehnt an medizinische Leitlinien, ausgedrückt im Grad einer Empfehlung. Dabei ist folgende Abstufung vorgesehen:

Empfehlung: Auf Basis der vorliegenden Erkenntnisse kann der Einsatz der DiGA für die medizinische Zielsetzung empfohlen werden.

Bedingte Empfehlung: Auf Basis der vorliegenden Erkenntnisse kann der Einsatz der DiGA für die medizinische Zielsetzung mit Einschränkungen empfohlen werden.

Keine Empfehlung: Auf Basis der vorliegenden Erkenntnisse kann der Einsatz der DiGA für die medizinische Zielsetzung derzeit nicht empfohlen werden.

Empfehlung offen: Auf Basis der vorliegenden Erkenntnisse kann derzeit noch keine Empfehlung bezüglich des Einsatzes der DiGA ausgesprochen werden.

Der Grad der Empfehlung resultiert aus dem qualitativen Gesamtbild der Gutachten und der Einordnungen durch die iAG – beispielsweise auch unter Einbeziehung des Umgangs mit möglichen unerwünschten Wirkungen. Dabei sind die Bewertungen für die Bereiche „Medizinisch-fachliche Fundierung“ und „Qualität der Evidenz der positiven Versorgungseffekte“ zwar nicht allein bedingend, aber richtungsweisend. Grundsätzlich lassen sich dazu folgende Regeln definieren:

- Für eine Empfehlung braucht es eine hohe medizinisch-fachliche Fundierung und eine hohe Qualität der Evidenz der positiven Versorgungseffekte.
- Eine Empfehlung mit Einschränkungen wird dann ausgesprochen, wenn für einen Bereich eine hohe Bewertung vergeben wurde und für den anderen eine mittlere.
- Keine Empfehlung wird ausgesprochen, wenn für beide Bereiche eine niedrige Bewertung vergeben wurde und auf Basis der vorliegenden Informationen nicht absehbar ist, dass sich das Gesamtbild in Kürze – etwa aufgrund neuer Studien – verändert.
- Wenn zu einem Bereich (noch) keine Daten vorliegen oder – zum Beispiel bei einer Kombination von mittlerer und niedriger Bewertung – noch nicht eindeutig absehbar ist, wie sich die medizinisch-fachliche Fundierung oder die Qualität der Evidenz der positiven Versorgungseffekte entwickelt, wird die Gesamtbewertung „Empfehlung offen“ ausgesprochen.

Mit der Einordnung „Empfehlung offen“ wird der Tatsache Rechnung getragen, dass es sich bei DiGA um eine vergleichsweise junge Technologie handelt und beispielsweise der Nachweis positiver Versorgungseffekte durch mehrere hochwertige Studien vielen Herstellern noch nicht möglich war. So sieht auch der Gesetzgeber für die Prüfung der Erstattungsfähigkeit durch

9 University of Oxford: Oxford Centre for Evidence-Based Medicine: Levels of Evidence, www.cebm.ox.ac.uk/resources/levels-of-evidence/oxford-centre-for-evidence-based-medicine-levels-of-evidence-march-2009 (abgerufen am 9.12.2021); Deutsches Netzwerk Evidenzbasierte Medizin e. V.: Arbeitsmaterialien, www.ebm-netzwerk.de/de/service-ressourcen/ebm-basics/arbeitsmaterialien (abgerufen am 9.12.2021); John Wiley & Sons: Essential Evidence PLUS. Levels of Evidence, https://www.essentialevidenceplus.com/product/ebm_loe.cfm?show=guidelines (abgerufen am 9.12.2021); Guyatt, G. H., et al. (2008): GRADE: an emerging consensus on rating quality of evidence and strength of recommendations, *BMJ*; 336:924.

TABELLE 5: Bewertungsschema „Medizinisch-fachliche Fundierung“

Medizinisch-fachliche Fundierung	Orientierungshilfe zur Bewertung
Hohe medizinisch-fachliche Fundierung	<p>Bezüge zwischen Zielsetzungen, Wirkweisen und / oder Nutzen der DiGA lassen sich aus Leitlinien und Publikationen eindeutig und plausibel herleiten.</p> <p>Die intendierten Wirkweisen und die dafür umgesetzten (Teil-)Funktionen der DiGA entsprechen voll und ganz dem aktuellen Stand des medizinischen Wissens.</p> <p>Die Patientensicherheit ist adäquat sichergestellt.</p>
Mittlere medizinisch-fachliche Fundierung	<p>Bezüge zwischen Zielsetzungen, Wirkweisen und / oder Nutzen der DiGA lassen sich aus Leitlinien und Publikationen mit Einschränkungen herleiten.</p> <p>Die intendierten Wirkweisen und die dafür umgesetzten (Teil-)Funktionen der DiGA entsprechen mit Einschränkungen dem aktuellen Stand des medizinischen Wissens.</p> <p>Die Patientensicherheit ist mit Einschränkungen sichergestellt.</p>
Niedrige medizinisch-fachliche Fundierung	<p>Bezüge zwischen Zielsetzungen, Wirkweisen und / oder Nutzen der DiGA lassen sich aus Leitlinien und Publikationen nicht ausreichend herleiten.</p> <p>Die intendierten Wirkweisen und die dafür umgesetzten (Teil-)Funktionen der DiGA entsprechen nicht ausreichend dem aktuellen Stand des medizinischen Wissens.</p> <p>Die Patientensicherheit ist nicht ausreichend sichergestellt.</p>
Keine Angabe / Datengrundlage	Auf Grundlage der verfügbaren Informationen kann die medizinisch-fachliche Fundierung nicht hinreichend beurteilt werden.

TABELLE 6: Bewertungsschema „Qualität der Evidenz der positiven Versorgungseffekte“

Qualität der Evidenz der positiven Versorgungseffekte	Orientierungshilfe zur Bewertung
Hohe Qualität der Evidenz der positiven Versorgungseffekte	<p>Der intendierte positive Versorgungseffekt ist durch die vorliegenden Studien gut nachgewiesen.</p> <p>Mehrere hochwertige Studien mit konsistenten Ergebnissen</p> <p>In besonderen Fällen: eine große, hochwertige multizentrische Studie</p>
Mittlere Qualität der Evidenz der positiven Versorgungseffekte	<p>Der intendierte positive Versorgungseffekt ist durch die vorliegenden Studien nachgewiesen. Das Evidenzniveau sollte durch weitere Studien erhöht werden.</p> <p>Eine hochwertige Studie</p> <p>Mehrere Studien mit einigen Einschränkungen</p>
Niedrige Qualität der Evidenz der positiven Versorgungseffekte	<p>Es ist weitere Forschung nötig, um den intendierten Versorgungseffekt nachzuweisen.</p> <p>Eine oder mehrere Studien mit starken Einschränkungen</p>
Noch keine Evidenz	Derzeit liegen noch keine Studien vor, um Aussagen zum Nachweis von positiven Versorgungseffekten zu treffen.

das BfArM im Rahmen des DiGA-Fast-Tracks ein Erprobungsjahr ohne vorliegende Primärstudien vor.

5.7 Aufbereitung und Kommunikation der Ergebnisse

Die Ergebnisse des Begutachtungsprozesses bereitet die wissenschaftliche Redaktion für die jeweilige DiGA in zwei Formaten auf: einem ausführlicheren Ergebnisbericht und einer kompakten Ergebniskarte, die sowohl Teil des Berichts ist als auch unabhängig davon veröffentlicht wird.

Der Ergebnisbericht beinhaltet unter anderem zentrale Informationen zu den für den Begutachtungsprozess verwendeten Daten, umfassendere Ausführungen zu den wesentlichen Teilen der Gutachten sowie möglichst praktische Hinweise für den Hersteller zu Möglichkeiten der Evidenzverbesserung.

Die Ergebniskarte besteht zum einen aus einer Kurzbeschreibung der DiGA sowie der medizinischen Zielsetzung und aus Hinweisen zur Anwendung (Indikationsbereich bzw. Zielgruppen). Die Informationen stammen vom Hersteller, werden aber redaktionell angepasst, um werbliche Inhalte zu vermeiden. Zum anderen werden die Bewertungsergebnisse für die Bereiche „Medizinisch-fachliche Fundierung“ und „Qualität der Evidenz der positiven Versorgungseffekte“ in knapper Form textlich vermittelt. Auch die Gesamtbewertung in Form eines Empfehlungsgrades ist Teil der Karte. Mit dem Ziel, dass die Nutzerinnen und Nutzer die Ergebnisse schnell erfassen können, werden die Empfehlungsgrade grafisch mit einem einfachen Star-Rating unterlegt:

Empfehlung: Drei von drei möglichen Sternen

Bedingte Empfehlung: Zwei von drei möglichen Sternen

Keine Empfehlung: Einer von drei möglichen Sternen

Empfehlung offen: Kein Star-Rating

Die Hersteller erhalten – da der Anspruch eines fairen Begutachtungsprozesses besteht – die Möglichkeit, die Ergebnisse zu ihrer DiGA zu kommentieren. Ein ausführlicher Kommentar wird im Ergebnisbericht veröffentlicht, eine kürzere Version auch in der Ergebniskarte. Zudem wird den Herstellern ein Gespräch zu den Ergebnissen angeboten (**Hersteller-Dialog**). Hierbei geht es um die Erläuterung der Ergebnisse, aber auch um einen Austausch zu den Möglichkeiten, die Qualität der Anwendung zu entwickeln, sowie zur Evidenzverbesserung.

Die Ergebniskarte und der Ergebnisbericht werden auf einer Website veröffentlicht, die von der Bertelsmann Stiftung bzw. der Weissen Liste betrieben wird. Zudem haben die Projektinitiatoren den Anspruch, die Ergebnisse möglichst breit in die Zielgruppe zu streuen, etwa, indem sie diese an einschlägige Fachgesellschaften verschicken.

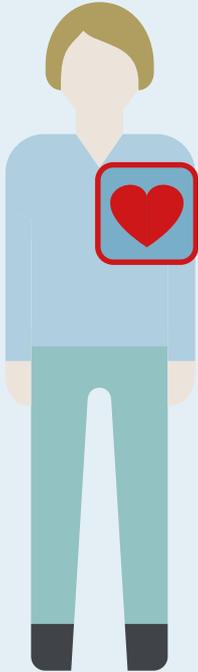
Die übergeordneten Erkenntnisse eines Bewertungsprojekts werden vom GBM, von der Bertelsmann Stiftung und der Weissen Liste ausgewertet und unter anderem in Form von Fachveröffentlichungen („Whitepaper“) in unregelmäßigen Abständen in den **Stakeholder-Dialog** eingebracht.

5.8 Aktualisierungen

Grundsätzlich gilt, dass die Bewertungsergebnisse zu einer DiGA für den Stand der finalen Datenerhebung gelten; dieser wird in den Ergebnisveröffentlichungen angegeben. Die Datenerhebung wird für alle DiGA innerhalb eines Bewertungsprojekts zu einem definierten Stichtag abgeschlossen. Anschließend sind aus Gründen der Gleichbehandlung aller Hersteller keine Aktualisierungen oder Ergänzungen mehr möglich. Sollten sich nach Abschluss eines Bewertungsprojekts und nach Ergebnisveröffentlichung wesentliche Aktualisierungen der Datengrundlagen ergeben, etwa durch Veröffentlichung einer neuen Primärstudie zur DiGA, werden diese von dem Team „Trusted-Health-Apps“ gemeinsam mit dem GBM-Vorsitz geprüft. Die Prüfung kann zu folgenden Resultaten und Reaktionen führen:

GUTACHTEN-BOARD MEDIZIN: GESUNDHEITS-APP

NAME DER ANWENDUNG (URL)



KURZ-BESCHREIBUNG

Was ist die Anwendung und wie wird die App eingesetzt?

...

MEDIZINISCHE ZIELSTELLUNG

Welche medizinische Zielstellung will die App erreichen?

...

ANWENDUNG

Bei welchen Indikationen soll die App eingesetzt werden?

...

GESAMT-BEWERTUNG

Ist der Einsatz der App für die medizinische Zielstellung aktuell empfehlenswert?

EMPFEHLUNG OFFEN

KEINE EMPFEHLUNG



MEDIZINISCH-FACHLICHE FUNDIERUNG

Inwieweit beruht die App auf dem aktuellen Stand des medizinischen Wissens?

...

QUALITÄT DER EVIDENZ

Wie hoch ist die Qualität der Evidenz für die positiven Versorgungseffekte der App?

...

KOMMENTAR DES HERSTELLERS

„...“

BEDINGTE EMPFEHLUNG



EMPFEHLUNG



Quellen: ...
 Datum der finalen Datenerhebung: ...
 Datum der Veröffentlichung: ...

WEISSE LISTE | BertelsmannStiftung

ABBILDUNG 2: Exemplarische Ergebniskarte

- Kein absehbarer Einfluss auf das Bewertungsergebnis: Veröffentlichung bleibt unverändert bestehen.
- Kein absehbarer Einfluss auf das Bewertungsergebnis, aber für die Nutzerinnen und Nutzer des Ergebnisses relevante Zusatzinformation: Veröffentlichung wird um einen Hinweis ergänzt.
- Absehbarer Einfluss auf das Bewertungsergebnis: (Temporäre) Rücknahme der Veröffentlichung und erneute Bewertung.

Die erneute Bewertung im Falle eines absehbaren Einflusses auf das Ergebnis kann auf einer weiteren iAG-Sitzung erfolgen oder – im Falle umfassender Aktualisierungen – in einem neuen Bewertungsprojekt zu einem späteren Zeitpunkt.

Wesentliche Aktualisierungen der Datengrundlagen können auf verschiedenen Wegen augenfällig werden, etwa durch Hinweise des Herstellers oder von Dritten oder durch Eigenrecherchen des Teams „Trusted-Health-Apps“. Dessen Anspruch ist es, die Aktualisierungen grundsätzlich innerhalb eines Zeitraums von acht Wochen nach Eingang eines Hinweises im Sinne des oben beschriebenen Prozesses zu prüfen, soweit dies für das Team im Einzelfall möglich und zumutbar ist. Bei einer Rücknahme der Veröffentlichung besteht kein Anspruch des Herstellers auf erneute Prüfung und/oder Veröffentlichung des Prüfergebnisses.

24

6 Rahmenbedingungen

Die in diesem Dokument beschriebene Methode dient dazu, einen Prozess zur medizinischen Begutachtung von DiGA zu etablieren, der zugleich fachlich fundiert und möglichst neutral ist. Um diese Ziele erreichen zu können, sind folgende Aspekte konstituierend für die Methode.

Definierter rechtlicher Rahmen: Das GBM und die iAGs werden auf Basis von Statuten tätig, die Rechte und Pflichten der Mitglieder festlegen und das Verhältnis zur Bertelsmann Stiftung sowie zur Weissen Liste transparent regeln. Die GBM-Mitglieder werden von der Stiftung und der Weissen Liste in der Regel für zwei Jahre berufen. Sie sind ehrenamtlich tätig, ausschließlich den Projektzielen sowie ihrer fachlichen Überzeugung verpflichtet und fungieren im Board nicht als Repräsentant oder Repräsentantin sonstiger Organisationen, denen sie angehören. Die Gutachterinnen und Gutachter erstellen ihre Gutachten auf Basis eines Vertrags, der die vorbezeichneten Anforderungen an ihre Tätigkeit absichern soll. Alle am Begutachtungsprozess Beteiligten erhalten für ihre Tätigkeit eine Aufwandsentschädigung, die aus Projektmitteln der Bertelsmann Stiftung finanziert wird.

Offenlegung von Interessenkonflikten: Mögliche Interessenkonflikte – im Rahmen eines Bewertungsprojekts – sollen bereits vor Durchführung eines Projekts vermieden und, soweit vorhanden, nach dessen Durchführung umfassend offenlegt werden. Das gilt sowohl für die Mitglieder des Boards als auch für die temporär tätigen iAG-Mitglieder sowie die externen Gutachterinnen und Gutachter. Die Offenlegung erfolgt spätestens mit Veröffentlichung der Ergebnisse eines Bewertungsprojekts.

Methodentransparenz: Der aktuelle Stand der Methode wird jeweils veröffentlicht, sodass für die Öffentlichkeit stets nachvollziehbar ist, auf welchem Weg die Ergebnisse einer Begutachtung zustande gekommen sind. Die Methode wird im kontinuierlichen Dialog mit der Fachöffentlichkeit und den Anspruchsgruppen des Projekts weiterentwickelt.

7 Ausblick

Die Ergebnisse der Begutachtung im GBM schaffen mehr Transparenz zur Qualität von DiGA. Sie widmen sich außerdem der neuen Herausforderung, digitale Anwendungen medizinisch zu bewerten. Der hier beschriebene Prozess wird kontinuierlich evaluiert und inhaltlich weiterentwickelt. Das Projekt soll so zu einer nachhaltigen Qualitätsorientierung im Feld und zu einer echten Verbesserung der Versorgung durch DiGA beitragen.

Die Bertelsmann Stiftung und die Weisse Liste streben eine nachhaltige Etablierung des Begutachtungsprozesses an. Perspektivisch ist ein Projekttransfer an einen öffentlichen bzw. gemeinwohlorientierten Akteur vorgesehen.

Impressum

© März 2022
Bertelsmann Stiftung

Bertelsmann Stiftung
Carl-Bertelsmannstraße 256
33311 Gütersloh
Telefon +49 5241 81-0
www.bertelsmann-stiftung.de

Verantwortlich
Uwe Schwenk

Autorinnen und Autoren
Prof. Dr. Sebastian Kuhn,
Timo Thranberend

Sina Busse,
Stefanie Hennig,
Jun.-Prof. Dr. Maike Henningsen,
Jun.-Prof. Dr. Stefanie Jungmann,
Prof. Dr. Jochen Klucken,
Dr. Johanna Ludwig,
Sabine Maur,
Prof. Dr. Martin Möckel,
Dr. Olaf Schmidt,
Prof. Dr. Christoph Schöbel,
Dr. Robin Weidemann

Lektorat
Heike Herrberg

Design
Dietlind Ehlers

Titelbild
© everythingpossible –
stock.adobe.com und
© alice_photo –
stock.adobe.com

Adresse | Kontakt

Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh
Telefon +49 5241 81-0

Sina Busse
Project Manager
Weisse Liste gemeinnützige GmbH
sina.busse@weisse-liste.de

Dr. Stefan Etgeton
Senior Expert
Programm „Gesundheit“
Bertelsmann Stiftung
stefan.etgeton@bertelsmann-stiftung.de
www.bertelsmann-stiftung.de

www.bertelsmann-stiftung.de