

Prozess und Bewertung digitaler Gesundheits- Anwendungen



am Beispiel der „Symptom-Checker“



HTA Austria
Austrian Institute for
Health Technology Assessment
GmbH

Prozess und Bewertung digitaler Gesundheits- Anwendungen

am Beispiel der „Symptom-Checker“

Projektteam

Projektleitung: Dr. med. Reinhard Jeindl

Projektbearbeitung: Dr. med. Reinhard Jeindl
Gregor Goetz, MSc MPH

Projektbeteiligung

Systematische Literatursuche: Tarquin Mittermayr, BA(Hons), MA

Visualisierungen: DI Smiljana Blagojevic

Externe Begutachtung: DI Dr. Robert Mischak, MPH
Dr. Dimitra Panteli, MScPH

Interne Begutachtung: Priv. Doz. Dr.phil. Claudia Wild

Korrespondenz: Reinhard Jeindl, reinhard.jeindl@aihta.at

Umschlagfoto: © elenabsl – stock.adobe.com

Dieser Bericht soll folgendermaßen zitiert werden/This report should be referenced as follows:

Jeindl R., Goetz G. Prozess und Bewertung digitaler Gesundheitsanwendungen – am Beispiel der „Symptom-Checker“. AIHTA Projektbericht Nr. 141, 2021. Wien: HTA Austria – Austrian Institute for Health Technology Assessment GmbH.

Interessenskonflikt

Alle beteiligten AutorInnen erklären, dass keine Interessenskonflikte im Sinne der Uniform Requirements of Manuscripts Statement of Medical Journal Editors (www.icmje.org) bestehen.

Disclaimer

The external reviewers did not co-author the scientific report and do not necessarily all agree with its content. Only the AIHTA is responsible for errors or omissions that could persist. The final version and the policy Recommendations are under the full responsibility of the AIHTA.

IMPRESSUM

Medieninhaber und Herausgeber:

HTA Austria – Austrian Institute for Health Technology Assessment GmbH
Garnisongasse 7/Top20 | 1090 Wien – Österreich
<https://www.aihta.at/>

Für den Inhalt verantwortlich:

Priv.-Doz. Dr. phil. Claudia Wild, Geschäftsführung

Die **AIHTA Projektberichte** erscheinen unregelmäßig und dienen der Veröffentlichung der Forschungsergebnisse des Austrian Institute for Health Technology Assessment.

Die **AIHTA Projektberichte** erscheinen in geringer Auflage im Druck und werden über den Dokumentenserver „http://eprints.aihta.at/view/types/hta_report.html“, der Öffentlichkeit zur Verfügung gestellt.

AIHTA Projektbericht Nr.: 141

ISSN 1993-0488

ISSN online 1993-0496

© 2021 AIHTA – Alle Rechte vorbehalten

Inhalt

Zusammenfassung	9
Executive Summary	13
1 Hintergrund	17
1.1 Einleitung, Definitionen	17
1.2 Abgestufter Bewertungsprozess	18
1.3 Symptom-Checker	20
2 Methodisches Vorgehen	21
2.1 Projektziel und Forschungsfragen	21
2.2 Übersicht zu digitalen Gesundheitsanwendungen, Priorisierung und technologische Anforderungen	21
2.3 Systematische Übersicht zu „Symptom-Checkern“	23
2.3.1 Literatursuche	23
2.3.2 Analyse und Qualitätsbewertung	25
2.3.3 Datenextraktion und Synthese	26
2.3.4 Endpunkte der eingeschlossenen Studien	26
2.4 Pilotevaluierung des Prozesses anhand einer ausgewählten DiGA („Symptoma“)	27
2.5 Feinkonzeption des Prozesses	28
2.6 Qualitätssicherung	28
3 Ergebnisse	29
3.1 Übersicht zu digitalen Gesundheitsanwendungen und deren Kategorisierung wie Priorisierung	29
3.1.1 Kategorisierung von DiGAs	29
3.1.2 Erstellung eines Meta-Registers von DiGAs	32
3.1.3 Priorisierungskriterien für die Evaluation von DiGAs in Österreich	33
3.2 Technologiespezifische Anforderungen an digitale Gesundheitsanwendungen	36
3.2.1 Technologische Anforderung für ELGA-Kompatibilität	36
3.2.2 Überblick über technologische Aspekte der DiGAs mit ASVG-Relevanz	40
3.2.3 Technologie-spezifische Aspekte: Risikoklassen, Datenschutz, Algorithmen, technische Standards für Interoperabilität, ELGA-Kompatibilität	52
3.3 Systematische Übersichtsarbeit zu Symptom-Checkern	55
3.3.1 Charakteristika der eingeschlossenen Studien	55
3.3.2 Evidenz der eingeschlossenen Studien	60
3.4 Pilotevaluierung des Prozesses anhand einer ausgewählten DiGA („Symptoma“)	65
3.4.1 Kategorisierung von „Symptoma“ und Bestimmung der ASVG Relevanz	65
3.4.2 Überprüfung der technologischen Anforderungen	66
3.4.3 Überprüfung der Risikoklasse nach MDR und Abgleich mit Evidenzanforderungen	66
3.5 Prozesskonzeption für die Bewertung von digitalen Gesundheitsanwendungen	69
4 Diskussion	73
4.1 Kategorisierung, Priorisierungskriterien und technologiespezifische Anforderungen	73
4.2 Pilotevaluierung des Bewertungsprozesses für DiGAs	74
4.2.1 Evidenz der Symptom-Checker	74
4.2.2 Weitere Erwägungen zu Symptom-Checkern	75
4.2.3 Beurteilung der Evidenz	76
4.3 Limitationen der Arbeit	76
5 Schlussfolgerungen	79
6 Literatur	81

Anhang	87
Qualitätsbewertung des NIHR Berichts mit dem AMSTAR-II Assessment Tool.....	87
Extraktionstabellen	88
Risk-of-Bias Assessment für diagnostische Studien mit dem QUADAS-II Assessment Tool.....	103
Meta-Verzeichnis von digitalen Gesundheitsanwendungen	105
Suchstrategien.....	109

Abbildungsverzeichnis

Abbildung 1-1: Skizzierter Bewertungsprozess für digitale Gesundheitsanwendungen, aus den Vorarbeiten des mHealth Berichts des AIHTAs 2020	19
Abbildung 2-1: Darstellung des Auswahlprozesses (PRISMA Flow Diagram)	25
Abbildung 3-1: Priorisierung der für Österreich relevanten DiGAs.....	35
Abbildung 3-2: Übersicht der Kompatibilität von digitalen Gesundheitsanwendungen mit der elektronischen Gesundheitsakte ELGA, basierend auf der Expert*innenkonsultation mit der ELGA GmbH	39
Abbildung 3-3: Prozesskonzeption für die Erstattung von digitalen Gesundheitsanwendungen in Österreich.....	72

Tabellenverzeichnis

Tabelle 2-1: Einschlusskriterien für die systematische Literatursuche	23
Tabelle 2-2: Risikoklassen der digitalen Gesundheitsanwendungen	27
Tabelle 3-1: Kategorisierungsansatz von DiGAs übernommen von Vorarbeiten des iDIGA-Projektes.....	31
Tabelle 3-2: Übersicht zu den im Meta-Verzeichnis inkludierten Verzeichnissen und Auflistung, angelehnt an den AIHTA Projektbericht 2020.....	32
Tabelle 3-3: Kategorisierung von DiGAs: Herz-Kreislaufkrankungen.....	42
Tabelle 3-4: Kategorisierung von DiGAs: Endokrine, Ernährungs- und Stoffwechselerkrankungen	44
Tabelle 3-5: Kategorisierung von DiGAs: psychische Erkrankungen	47
Tabelle 3-6: Kategorisierung von DiGAs: andere Indikationen	51
Tabelle 3-7: Technologiespezifische Anforderungen und Einteilung der DiGAs in Risikoklassen sowie Evidenzstufen.....	53
Tabelle 3-8: Wesentliche Merkmale der systematischen Übersichtsarbeit des NIHR-Berichts	58
Tabelle 3-9: Wesentliche Merkmale des Studientyps, Studiensettings, verwendeten Interventionen und gemessenen Endpunkten der zusätzlich identifizierten Studien.....	59
Tabelle 3-10: Evidenzstufen für digitale Gesundheitsanwendungen	67
Tabelle 3-11: Technologiespezifische Charakteristika der DiGA „Symptoma“, basierend auf „Module for MMA“ und „Digi-HTA Framework“	68
Tabelle 3-12: Fragenkatalog zu den verfahrenstechnischen, offenen Fragen zum Bewertungsprozess für digitale Gesundheitsanwendungen	70
Tabelle A-1: Bewertung der Qualität der systematischen Übersichtsarbeit des National Institute of Health Research unter Anwendung des AMSTAR-II Bewertungsinstruments	87
Tabelle A-2: Datenextraktionstabelle zur systematischen Übersichtsarbeit des NIHR-Berichts	88
Tabelle A-3: Datenextraktionstabelle zu zusätzlich identifizierten Studien zu Symptom-Checkern (Teil 1).....	90
Tabelle A-3: Datenextraktionstabelle zu zusätzlich identifizierten Studien zu Symptom-Checkern (Teil 2).....	93
Tabelle A-4: Datenextraktionstabelle zu zusätzlich identifizierten Studien zur Patient*innensicht von Symptom-Checkern.....	97

Tabelle A-5: Einstufung des Verzerrungspotenzials (Risk-of-Bias) für diagnostische Studien unter Anwendung des QUADAS-II Bewertungsinstruments..... 103

Tabelle A-6: Meta-Verzeichnis der deutschsprachig verfügbaren digitalen Gesundheitsanwendungen aus ausgewählten Verzeichnissen (kategorisiert nach Zielgruppe und Funktion) 105

Abkürzungsverzeichnis

Abs.....Absatz	FGFachgebiet
AIHTAAustrian Institute for Health Technology Assessment	GDAGesundheitsdienstanbieter
AITAustrian Institute of Technology	GDPR.....Allgemeine Datenschutz-Verordnung („General Data Protection Regulation“)
AMSTAR ...A Measurement Tool to Assess Systematic Reviews	GPAllgemeinmediziner*in („General practitioner“)
Art.Artikel	HHV.....Health Hub Vienna
ASVG.....Allgemeines Sozialversicherungs-Gesetz	HL7Health Level Seven International
AT.....Österreich	HTAHealth Technology Assessment
BE.....Belgien	ICD-10International Classification of Diseases, 10. Ausgabe
BfArMBundesinstitut für Arzneimittel und Medizinprodukte (Deutschland)	IfGP.....Institut für Gesundheitsförderung und Prävention
CDAClinical Document Architecture	INAHTA...International Network of Agencies for Health Technology Assessment
CDDS.....Computerised diagnostic decision support programmes	INiTSUniversitäres Gründerservice Wien
CDSS.....Computerised clinical decision support systems	ITSV.....IT-Services der Sozialversicherung
CE.....europäische Konformität („Conformité européenne“)	k.A.keine Angabe
CI.....Konfidenzintervall	KISKrankenhaus-informationssystem
CRDCentre for Reviews and Dissemination	LE.....Leistungserbringer
DEDeutschland	LPPR.....Regelwerk für Erstattungsentscheidungen zu Medizinprodukten („Liste des Produits et Prestations Remboursables“, Frankreich)
DiGA.....Digitale Gesundheitsanwendung	MDRMedizinprodukt-verordnung („Medical device regulation“)
DiGAsDigitale Gesundheitsanwendungen	mHealthmobile Health
DVGDigitale-Versorgung-Gesetz (Deutschland)	MMA.....digitale Gesundheitsanwendung („Mobile Medical Application“)
DVSVDachverband der Sozialversicherungsträger	n.....Anzahl („number“)
EDNotaufnahme („Emergency department“)	NA.....nicht zutreffend („not applicable“)
EKO.....Erstattungskodex	NeLL.....National eHealth Living Lab (Niederlande)
ELGAElektronische Gesundheitsakte	NHSNational Health Service (Großbritannien)
EUEuropäische Union	NICE.....National Institute for Health and Care Excellence (Großbritannien)
FFForschungsfrage	
FHIR.....Fast Healthcare Interoperability Resources	

NIHR	National Institute for Health Research (Großbritannien)	SD.....Standardabweichung
Nr.	Nummer	SoC.....herkömmliche Versorgung („Standard of Care“)
NRCT	nicht-randomisiert kontrollierte Studie („non-randomized controlled trial“)	SR.....Systematische Übersichtsarbeit („Systematic review“)
p	p-Wert	QUADAS ...Quality Assessment of Diagnostic Accuracy Studies
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses	UK.....Großbritannien
RCT	randomisiert kontrollierte Studie („randomized controlled trial“)	USA.....Vereinigte Staaten von Amerika
RoB	Verzerrungspotenzial („Risk of Bias“)	WHO.....World Health Organization
		Z-PI.....zentraler Patient*innen-Index

Zusammenfassung

Hintergrund

Digitale Gesundheitsanwendungen (DiGAs) werden zunehmend mit diagnostischen und therapeutischen Funktionen angeboten. Häufig ist der medizinische oder organisatorische Zusatznutzen jedoch unklar. Im internationalen Kontext stellt sich demnach für Krankenkassen und Sozialversicherungen die Frage, wie eine evidenzbasierte Erstattung der DiGAs in die Regelversorgung gestaltet werden könnte.

In Österreich gibt es bislang noch keinen etablierten Prozess für den Umgang mit DiGAs. Im Jahr 2020 publizierte das AIHTA einen Bericht „Framework zur Unterstützung von Refundierungsentscheidungen zu digitalen Gesundheitsanwendungen“, in welchem international verfügbare Bewertungsinstrumente für DiGAs analysiert wurden. Ziel des vorliegenden Berichts war es, eine weiter verfeinerte, praxistaugliche Prozesskonzeption für die Erstattung von DiGAs in Österreich vorzuschlagen.

Dazu wurde der skizzierte Prozess des AIHTA-Berichts 2020 nun an einer DiGA aus der Gruppe der sogenannten „Symptom-Checker“ pilotiert. Diese Symptom-Checker erstellen anhand eingegebener Symptome eine Liste mit möglichen Ursachen, und sollen dadurch den Diagnoseprozess unterstützen und für eine verbesserte Steuerung von Patient*innenflüssen sorgen.

Methoden

Meta-Verzeichnis und Kategorisierung von DiGAs

Zunächst wurde ein Meta-Verzeichnis von DiGAs erstellt. Dabei wurden DiGAs von Listungen und Verzeichnissen aus vier Ländern (Belgien, Deutschland, Vereinigtes Königreich und Österreich) eingeschlossen. Für die Identifikation von DiGAs aus Österreich, der Erstellung von Priorisierungskriterien, und der technologiespezifischen Anforderungen von DiGAs wurden Expert*innenkonsultationen durchgeführt. Die Konsultationen wurden mit Expert*innen des Allgemeinen Sozialversicherungsgesetzes (ASVG), der elektronischen Gesundheitsakte (ELGA), den IT-Services der Sozialversicherungen (ITSV), und mit Expert*innen mit Kenntnissen der DiGAs am österreichischen Markt (Dachverband der Sozialversicherungsträger, Health Hub Vienna, Austrian Institute of Technology) durchgeführt. Basierend auf den von der Technischen Universität Berlin erstellten Kategorisierungsansätzen für DiGAs wurden die DiGAs des Meta-Verzeichnisses in entsprechende Kategorien eingeteilt.

Systematische Übersichtsarbeit zu Symptom-Checkern

Zur anschließenden Analyse der Evidenz zu Symptom-Checkern wurde im vorliegenden Bericht eine britische Evidenzsynthese des National Institute for Health Research (NIHR) zu Symptom-Checkern (2019) aktualisiert. Die methodische Qualität der systematischen Übersichtsarbeit des NIHR-Berichts wurde anhand der AMSTAR-II Checkliste bewertet.

Es wurde eine Update-Suche der Evidenz anhand der Suchstrategien des NIHR-Berichts durchgeführt. Dafür wurden vier medizinischen Datenbanken im Zeitraum ab 2018 für die systematische Suche herangezogen. Die Studienauswahl, Datenextraktion und Bewertung der methodischen Qualität der

Konzept einer evidenzbasierten Erstattung von DiGAs

Österreich noch ohne etablierten Prozess

Pilotevaluierung anhand eines „Symptom-Checkers“

Meta-Verzeichnis europäischer DiGAs

Konsultationen mit juristischen und technologischen Expert*innen

update HTA: National Institute for Health Research Bericht (2019)

Datenextraktion und Bewertung von 2 Personen durchgeführt

Studien wurden von zwei Personen unabhängig voneinander durchgeführt. Nach der Datenextraktion folgte eine narrative Beschreibung der Evidenz.

**ausgewählte
Endpunkte**

Für die Bewertung wurden folgende Endpunkte gewählt:

- Wirksamkeit: Krankheitsdauer und Schwere der Erkrankung,
- Sicherheit: Schäden durch falsch-positive/falsch-negative Diagnosevorschläge oder Triage-Empfehlungen,
- Diagnostische Genauigkeit,
- Genauigkeit der Triage-Empfehlungen,
- Systemisch/organisatorische Endpunkte: Anzahl der Arzt-/Ambulanzbesuche,
- Patient*innenrelevante Endpunkte: Lebensqualität und Zufriedenheit der Patient*innen.

Prozesskonzeption für Refundierung von DiGAs

**Feinkonzeption des
Bewertungsprozesses für
DiGA-Refundierung**

Basierend auf den Erkenntnissen der Expert*innenkonsultationen, und der Pilotevaluierung anhand der Symptom-Checker, wurde abschließend der skizzierte Prozess aus dem AIHTA Projektbericht 2020 in einer Feinkonzeption um weitere Bausteine im Prozess für Refundierungsentscheidungen zu DiGAs ergänzt.

Ergebnisse

Meta-Verzeichnis und Kategorisierung von DiGAs

**132 DiGAs im
Meta-Verzeichnis**

Das Meta-Verzeichnis von DiGAs umfasst 132 mobile App- und Webanwendungen. Nach dem Kategorisierungsansatz der Technischen Universität Berlin können die DiGAs anhand von vier Kategorien eingeteilt werden: Anwendungsbereich (nach ICD-10 Code), Anwender*in (Patient*in und/oder Leistungserbringer), Zielgruppe (Vulnerabilität der Patient*in) sowie Funktion der DiGA (Erkennung/Diagnose, Überwachung/Monitoring und Behandlung/Therapie).

**Kategorienansatz der
TU Berlin**

**Priorisierungskriterien
nach ASVG:
derzeit von 38 DiGAs erfüllt**

Aufbauend auf den Bestimmungen des ASVG wurden folgende Priorisierungskriterien entwickelt:

- Anwendung: in Deutsch verfügbar
- Einschränkung auf Chronikerprogramme
- Funktionen: Behandlung und Monitoring

**unterschiedliche
Indikationsbereiche**

Von den in Deutsch verfügbaren DiGAs erfüllen derzeit 38 diese Priorisierungskriterien. Die Indikationsbereiche der priorisierten DiGAs umfassen psychische Störungen (n=9), Endokrine, Ernährungs- und Stoffwechselerkrankungen (n=7), Herz-Kreislaufkrankungen (n=7) und „andere“ Indikationen (n=15).

**CE-Kennzeichnung,
Datenschutz, Angaben
zur Interoperabilität**

Die Überprüfung der technologiespezifischen Charakteristika zeigte, dass der Großteil der DiGAs über eine CE-Kennzeichnung (n=29) und Angaben zum Datenschutz (n=35) verfügt. Die Kompatibilität der DiGAs mit der elektronischen Gesundheitsakte (ELGA) ist derzeit noch nicht ausgereift. Von 14 DiGAs mit Angaben zu verwendeten Interoperabilitäts-Standards erfüllen derzeit sieben DiGAs die Voraussetzungen für eine technische Kompatibilität mit ELGA. Die Verwendung des „HL7 CDA“ Standards der DiGAs lässt jedoch noch nicht auf eine semantische Kompatibilität mit ELGA schließen.

Systematische Übersichtsarbeit zu Symptom-Checkern

Die Evidenz der NIHR-Übersichtsarbeit über Symptom-Checker umfasste 27 Studien und wurde durch 14 weitere Studien aktualisiert.

Sicherheit: In den Studien wurden keine möglichen Schäden durch Verwendung der Symptom-Checker identifiziert (sechs Beobachtungsstudien).

Wirksamkeit: Bei Verwendung von Symptom-Checkern gab es eine geringfügig längere Krankheitsdauer, statistisch nicht signifikant (ein RCT).

Die Evidenz zur *Genauigkeit der Diagnosevorschläge* und der *Triage-Empfehlungen* lieferte keine zufriedenstellenden Ergebnisse. Die Symptom-Checker wurden sehr häufig anhand von fiktiven, klinischen Fallvignetten im Rollenspielszenario getestet. Da nicht ausreichend ausgeschlossen werden kann, dass die Algorithmen der Symptom-Checker mit denselben Fallvignetten trainiert wurden, welche für die Erhebung herangezogen wurden, ist eine Übertragbarkeit auf reale Szenarien sehr eingeschränkt. Für die *diagnostische Genauigkeit* reichte die Spannweite von 14-84,3 % (zehn Beobachtungsstudien), die *Genauigkeit der Triage-Einschätzungen* reichte von 33-100 % (elf Beobachtungsstudien).

Für die *systemisch/organisatorischen Endpunkte* kann aufgrund von inkonsistenter Evidenz keine Aussage über eine Zu- oder Abnahme von Arzt-/Ambulanzbesuchen gemacht werden (ein RCT und sechs Beobachtungsstudien).

Die neun qualitativen Survey-Studien zur *Patient*innensicht* deuten darauf hin, dass die meisten Nutzer*innen bislang keine Erfahrung mit der Verwendung von Symptom-Checkern haben. Das Symptom-Assessment wurde als strukturierter als eine herkömmliche Internet-Recherche wahrgenommen, wenngleich die beschränkte Möglichkeit der vollständigen Beschreibung von Symptomen und das Fehlen der verbalen Interaktion mit dem Gesundheitspersonal als hinderlich angegeben wurden.

Prozesskonzeption für Refundierung von DiGAs

Die Prozessschritte aus dem AIHTA-Bericht 2020 „Framework zur Unterstützung von Refundierungsentscheidungen zu digitalen Gesundheitsanwendungen“ wurden nach der Pilotevaluierung und den Expert*innenkonsultationen adaptiert. Diese umfassen nun folgende Prozessschritte:

- *Identifikation:* Antragsverfahren für DiGAs, oder DiGAs welche in Kooperation mit den Sozialversicherungen entwickelt wurden,
- *Filterprozesse:* Art der Anwendung, Sprache, Priorisierungskriterien nach Relevanz entsprechend dem ASVG,
- *Anforderungen:* Datenschutz, Interoperabilität, regulatorische Anforderungen, Evidenzanforderungen,
- *Evaluation:* Health Technology Assessment zur Nutzenbewertung,
- *Refundierung:* bei erbrachtem Nutznachweis, geknüpft an Kriterien für Re-Assessment der DiGA

27 Studien (NIHR) durch 14 zusätzliche Studien ergänzt

Ergebnisse nicht zufriedenstellend, Nutznachweis nicht erbracht

Mängel in Studien (klinische Fallvignetten, Rollenspielszenarien)

hohe Anwendbarkeit, aber fehlende verbale Interaktion; vollständige Anamnese nicht möglich

Adaptierung der Prozessschritte

Diskussion und Schlussfolgerung

Zusammenarbeit österreichischer Institutionen empfohlen

In der gründlichen Bewertung von DiGAs ist die Zusammenarbeit verschiedener österreichischer Institutionen mit entsprechender Aufgabenverteilung empfehlenswert (insbesondere für die Prüfung der Datenschutzvorgaben und der Kompatibilität mit ELGA).

ASVG Fokus: Behandlung/Monitoring von chronisch Erkrankten

Die Priorisierungskriterien richten sich nach den Bestimmungen des ASVG (Fokus auf Anwendungen für chronisch Erkrankte mit Funktionen zur Behandlung und Monitoring). Gegebenenfalls können die technischen Voraussetzungen für eine Kompatibilität mit ELGA ein weiteres Filterkriterium vor einer Evaluierung darstellen.

Symptom-Checker: diagnostische DiGAs, Nutznachweis nicht erbracht

DiGAs aus der Gruppe der Symptom-Checker erfüllen die ASVG-relevanten Priorisierungskriterien nicht stringent, da sie der Funktionsgruppe „Erkennung/Diagnose“ zuzuordnen sind. Für Symptom-Checker konnte der Nutznachweis nicht ausreichend erbracht werden.

Klärung noch offener verfahrenstechnischer Fragen erforderlich

Der in diesem Bericht vorgeschlagene Prozess für die Refundierung von DiGAs stellt eine wissenschaftliche Unterstützung für politische Entscheidungen dar; dieser kann nach Klärung der noch offenen verfahrenstechnischen Fragen weiter adaptiert werden.

Executive Summary

Background

Digital health applications (“DiGAs”) are increasingly offered with diagnostic and therapeutic functionalities. However, the medical or organizational additional benefit is often unclear. Internationally, the process of reimbursing these DiGAs in a standardized way is discussed by health and social insurances.

In Austria, there is not yet an established process for handling DiGAs. In 2020, the AIHTA published a report “Framework for reimbursement decisions of digital health technologies (mHealth)”. The aim of the current report was to further adapt the process outline and recommend a granular and practical design for a reimbursement process of DiGAs in Austria.

The process outline from the AIHTA-report 2020 was tested on a DiGA from the group of the so-called “Symptom-checkers”. These symptom-checkers claim to improve the diagnostic process and guidance of patients by offering a list of potential diseases, based on entered symptoms.

**conceptualization
of an evidence-based
reimbursement of DiGAs**

**Austria: not yet an
established process
for DiGAs**

**trial assessment on
“symptom-checkers”**

Methods

Meta-directory and categorization of DiGAs

At first, a meta-directory of DiGAs based on lists from four countries (Belgium, Germany, United Kingdom and Austria) was created. For the identification of lists from Austria, as well as categorization and prioritization criteria and technology-specific requirements for DiGAs, consultations with experts of the Austrian law on general health insurance (ASVG), the Austrian electronic health record and IT-services of the social insurances (ELGA, ITSV) as well as experts with knowledge of DiGAs on the Austrian market were conducted. Based on categorization criteria for DiGAs (developed by the Technical University of Berlin), the DiGAs of the meta-directory were classified.

**meta-directory
of European DiGAs**

**consultations with legal
and technological experts**

Systematic review on symptom-checkers

Thereafter, for an analysis of the evidence on symptom-checkers, a British evidence synthesis from the National Institute for Health Research (NIHR) on symptom-checkers (2019) was updated. The methodological quality was assessed using the AMSTAR-II checklist.

**update HTA:
National Institute for
Health Research (2019)**

An update search for new evidence was conducted. Four databases were selected for the systematic search. Additionally, a hand search was performed. After deduplication, a total of 1,009 potential hits were identified. The study selection, data extraction and assessment of methodological quality of the studies were independently performed by two researchers. After the data extraction, a narrative description of the evidence was presented.

**data extraction
and methodological
assessment by
2 independent researchers**

For the assessment, the following endpoints were selected: effectiveness, safety, diagnostic accuracy, accuracy of triage disposition, systemic/organizational endpoints, as well as patient relevant endpoints.

selected endpoints

<p>further conceptualization of reimbursement process</p>	<p>Design of a process for the reimbursement of DiGAs</p> <p>Based on the insights from the expert consultations as well as the assessment of symptom-checkers, the process outline from the AIHTA report 2020 was further adapted with additional suggested steps in the reimbursement decision process for DiGAs.</p>
<p>meta-directory: 132 DiGAs</p> <p>categorization system by TU Berlin</p>	<p>Results</p> <p>Meta-directory and categorization of DiGAs</p> <p>The meta-directory of DiGAs contains 132 mobile apps and browser applications. Considering the legal implications of the ASVG, prioritization criteria were developed (application available in German language, focus on therapeutic and monitoring functions for chronic disease management). Currently, 38 DiGAs fulfill these prioritization criteria. The indications of these prioritized DiGAs span over mental illness (n=9), endocrinological, nutritional and metabolic diseases (n=7), cardiovascular diseases (n=7) and „other“ indications (n=15).</p>
<p>prioritization criteria based on ASVG: fulfilled by 38 DiGAs</p>	<p>When assessing the technology-specific considerations, for the majority of DiGAs a CE-certificate (n=29) and a statement on data protection (n=35) was available. The interoperability of the DiGAs with the electronic health record (ELGA) is not yet fully integrated. Out of 14 DiGAs with a description on utilized technical standards for interoperability, seven DiGAs fulfill the requirements for a technical compatibility with ELGA. However, the utilization of the „HL7 CDA“ technical standard from a DiGA does not ensure a semantic compatibility with ELGA.</p>
<p>27 studies (NIHR), updated with 14 additional studies</p>	<p>Systematic review on symptom-checkers</p> <p>The evidence of the NIHR-report on symptom-checkers included 27 studies and was updated with further 14 studies.</p> <p><i>Safety:</i> In the studies, no possible harms when using symptom-checkers were identified (six observational studies).</p> <p><i>Effectiveness:</i> a slightly longer duration of illness when using symptom-checkers was reported, statistically not significant (one RCT).</p>
<p>results not satisfactory, evidence insufficient to show medical or organizational benefit</p> <p>studies with flaws (fictional case vignettes, role-play scenarios)</p>	<p>The evidence for <i>diagnostic accuracy</i> and <i>accuracy of triage disposition</i> didn't provide satisfactory results. The symptom-checkers were often assessed using fictional clinical case vignettes in role-play scenarios. It is possible that the algorithms of the symptom-checkers were trained with the same case vignettes that were used in the studies, therefore the applicability to real scenarios is very limited. The <i>diagnostic accuracy</i> ranged from 14-84.3% (ten observational studies), the <i>triage accuracy</i> ranged from 33-100% (eleven observational studies).</p> <p>For the systemic/organizational endpoints (increase or decrease of doctor or outpatient visits), a comprehensive statement is not yet possible, due to inconsistent results (one RCT and six observational studies).</p>
<p>usability high, but verbal interaction with health personnel lacking</p>	<p>The nine qualitative survey-studies on the patient perspective indicate, that the majority of users so far have no experience with using symptom-checkers. The symptom assessment was perceived as more structured compared to a conventional web search. However, the limited options for a full description of symptoms and the missing verbal interaction with health personnel were mentioned as hindering factors.</p>

Design of a process for the reimbursement of DiGAs

The suggested outline for the process was adapted to include process steps for the identification (application process for DiGAs, or DiGAs that were created in cooperation with health insurances), filter processes (type of application, language, prioritization criteria based on the ASVG), requirements (data protection, interoperability, regulatory requirements, evidence requirements) and evaluation (health technology assessment). For DiGAs with sufficient evidence for benefit, a reimbursement can be implemented. The reimbursement is also linked to criteria for a re-assessment.

**adaptation
of the process design**

Discussion and Conclusion

In the thorough evaluation of DiGAs, the cooperation of different Austrian institutes with corresponding task distribution is recommended (in particular for the assessment of data protection and interoperability with ELGA).

**cooperation of different
institutes recommended**

The prioritization criteria are based on the legal conditions of the ASVG (focus on applications with therapeutic and monitoring functionalities for chronic diseases). The technical requirements for an interoperability with ELGA could potentially pose another filter criterion before evaluating a DiGA.

**ASVG: focus on chronic
disease management
programmes**

DiGAs from the group of symptom-checkers don't fulfil the ASVG prioritization criteria, because they belong to the category of „detection/diagnosis“. For symptom-checkers, the currently available evidence is insufficient to show a medical or organizational benefit.

**symptom-checkers:
insufficient evidence
for diagnostic functions**

The designed process for the reimbursement of DiGAs of this report offers a scientific guidance for political decisions. After clarifying certain procedural and policy questions, the process for the reimbursement of DiGAs can be further adapted.

**need for addressing
procedural questions**

1 Hintergrund

1.1 Einleitung, Definitionen

mHealth („mobile Health“) ist als Komponente von eHealth definiert und bedeutet den Einsatz von mobilen Geräten wie Mobiltelefonen und anderen drahtlosen Geräten zum Monitoring von Vitalwerten von Patient*innen im medizinischen Kontext bzw. im Kontext der (öffentlichen) Gesundheitsversorgung [1]. Ein zunehmend häufiger verwendeter, etwas breiterer Begriff ist die Bezeichnung der digitalen Gesundheitsanwendungen (DiGAs). Diese sind definiert als „digitale Medizinprodukte, die die Versicherte oder den Versicherten etwa bei der Behandlung von Erkrankungen oder dem Ausgleich von Beeinträchtigungen unterstützen können. Erscheinungsformen sind neben Gesundheits-Apps etwa browserbasierte Webanwendungen oder Software zur Verwendung auf klassischen Desktop-Rechnern [2].“

Im internationalen Kontext spielen DiGAs für öffentliche Zahler zunehmend eine Rolle: bisher werden DiGAs großteils auf Basis von gesonderten Vereinbarungen (z. B. von einzelnen Krankenkassen in Selektivverträgen) erstattet. Erstattungswege in die Regelversorgung sind aber bisher in den meisten Ländern nicht vorhanden. Inzwischen gibt es international Bestrebungen mit Bewertungs- und Refundierungsmodellen für eine verbesserte Regulierung von medizinischer Software zu sorgen[3].

In Deutschland etwa wurde mit dem Digitale-Versorgung-Gesetz (DVG, Dez. 2019) [4] das Verschreiben und die Refundierung von ausgewählten DiGAs durch Ärzt*innen und Therapeut*innen beschlossen – zu diesem Zweck wurde durch das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) ein Verzeichnis erstattungsfähiger digitaler Gesundheitsanwendungen (DiGA-Verzeichnis) erstellt. Derzeit befinden sich 22 erstattungsfähige DiGAs (Stand: 10.10.2021) in diesem Verzeichnis [5].

In Belgien wurde der administrative Rahmen für die Regulierung von DiGA durch die Plattform „mHealthBelgium“ bereits geschaffen – für DiGAs mit nachgewiesenem sozioökonomischem Nutzen ist eine Refundierung der öffentlichen Krankenversicherungen geplant. In Frankreich werden DiGAs im bestehenden Regelwerk für Erstattungsentscheidungen zu Medizinprodukten (LPPR/„Liste des Produits et Prestations Remboursables“) eingegliedert. In den Niederlanden besteht mit der Einrichtung des National eHealth Living Lab (NeLL) eine nationale App-Listung. In Großbritannien gibt es mit der National Health Service (NHS) Apps Library eine Listung aller (nachgewiesenen) DiGAs, jedoch ohne Rückschlüsse zu Erstattungsfähigkeit. In Österreich gibt es bislang noch keinen etablierten Prozess für den Umgang mit DiGAs.

Die Detailbeschreibung der einzelnen Länderstrategien und die Empfehlungen für Österreich findet sich im Projektbericht 2020 „Framework zur Unterstützung von Refundierungsentscheidungen zu digitalen Gesundheitsanwendungen (mHealth) und dessen (retrospektive) Anwendung an ausgewählten Beispielen“ [6].

Auf regulatorischer Ebene gibt es, durch die EU-Medizinproduktverordnung 2017/745 (seit Mai 2021 in Kraft getreten), nennenswerte Entwicklungen in Bezug auf DiGAs. Durch diese Verordnung ist medizinische Software als Me-

**mHealth/DiGA:
digitale Medizinprodukte**

**internationale
Entwicklungen:
Regulierung und
Refundierung**

**DE: DiGA-Verzeichnis
(22 DiGAs)**

**Rahmenbedingungen
für DiGAs in Belgien,
Frankreich, Großbritannien**

**Österreich noch ohne
etablierten Prozess**

**regulatorische
Entwicklungen:
EU MDR 2017/745**

<p>hohe Erwartungen, bisher wenig Evidenz</p>	<p>dizinprodukt einzustufen, wenn die Software laut Hersteller dazu bestimmt ist, medizinische Zwecke zu erfüllen. In Folge dieser neuen EU-Verordnung 2017/745 ist es zu einer Höherstufung der Risikoklassen vieler DiGAs gekommen [6].</p> <p>Die Erwartungen an die Entlastung des Gesundheitssystems durch diese digitalen Gesundheitsanwendungen (DiGAs) sind hoch, wenngleich für die Mehrzahl der verfügbaren DiGAs wenig Evidenz zum tatsächlichen Nutzen vorliegt. Aufgrund der iterativen Entwicklung und Software-Updates von DiGAs ergeben sich entsprechende Herausforderungen in der Evaluation [7]. Die gründliche Evaluation von DiGAs, welche erstattet resp. bezuschusst werden, erfordert ein nachvollziehbares methodisches Vorgehen.</p>
<p>Teil 1: abgestufter Bewertungsprozess</p>	<p>Im vorausgegangenen Projektbericht 2020 zu digitalen Gesundheitsanwendungen [6] wurde, basierend auf einer systematischen Literatursuche nach Bewertungsinstrumenten für DiGAs, ein abgestuftes Vorgehen zur Bewertung von DiGAs in Österreich empfohlen. Dieses ist in Kapitel 1.2 im Detail beschrieben.</p>
<p>Teil 2: Feinkonzeption des Prozesses</p>	<p>Bisher wurde das im Projektbericht 2020 [6] beschriebene methodische Konzept zur Evaluierung von digitalen Gesundheitsanwendungen nicht erprobt. In diesem Folgeprojekt soll ein Vorgehen konzipiert werden, wie dieser Bewertungsprozess konkret in den Erstattungsprozess eingebettet werden könnte und welche Voraussetzungen an DiGA gestellt werden, bevor diese in die Versorgung integriert werden.</p>
<p>Pilotierung anhand „Symptom-Checker“</p>	<p>Zur Pilotierung des Bewertungsframeworks soll eine DiGA aus der Gruppe der „Symptom-Checker“ anhand der vorliegenden Evidenz bewertet und mit ähnlichen/gleichen Anwendungen verglichen werden [8, 9]. Als Ergebnis soll eine verfeinerte praxistaugliche Prozesskonzeption vorliegen.</p>

1.2 Abgestufter Bewertungsprozess

**Bewertungsprozess:
Überprüfung, Abgleich,
Screening, HTA-Aspekte**

Der im Projektbericht 2020 „Framework zur Unterstützung von Refundierungsentscheidungen zu digitalen Gesundheitsanwendungen (mHealth) und dessen (retrospektive) Anwendung an ausgewählten Beispielen“ [6] empfohlene Bewertungsprozess sieht vor,

- in einem ersten Schritt die Überprüfung der CE-Kennzeichnung sowie eine Einteilung in Risikoklasse nach Medizinproduktverordnung EU MDR 2017/745 vorzunehmen [10].
- Anschließend soll die MDR-Risikoklasse mit den Evidenzanforderungen der Evidenzstufen des NICE Evidence Standards Frameworks abgeglichen werden [11]. Randomisierte, kontrollierte Studien, die die DiGA mit dem derzeitigen Versorgungsstandard vergleichen, stellen das höchste Maß an Evidenz dar.
- In einem nächsten Schritt sollte ein Screening der vorhandenen Evidenz erfolgen.
- DiGA ohne Studien, die den Evidenzanforderungen des NICE Evidence Standards Frameworks entsprechen, sollten (vorerst) von einer Refundierung ausgeschlossen werden, bis ausreichend Evidenz vorgelegt wird. Die jeweiligen Evidenzanforderungen je nach Evidenzstufe finden sich im AIHTA-Projektbericht 2020.

- Für DiGA, welche die Anforderungen erfüllen, können dann die relevanten HTA-Aspekte definiert, die sich an den Erwartungen der Krankenkassen und Versprechungen der Hersteller orientieren, und infolge bewertet werden, um eine fundierte Entscheidung zum nachgewiesenen Nutzen und den zu erwartenden Effekten der digitalen Gesundheitsanwendung treffen zu können. Hierfür empfiehlt sich eine Orientierung am HTA Module for Mobile Medical Applications [12] und dem Digi-HTA Framework [13].

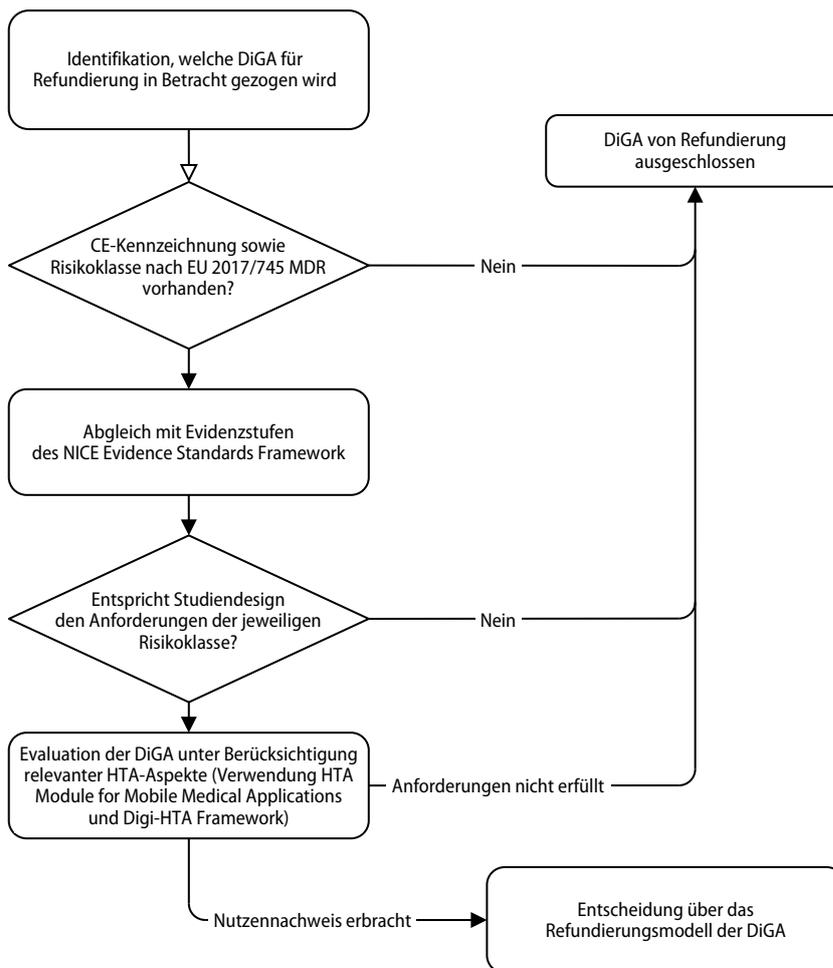


Abbildung 1-1: Skizzierter Bewertungsprozess für digitale Gesundheitsanwendungen, aus den Vorarbeiten des mHealth Berichts des AIHTAs 2020 [1]

Im skizzierten Bewertungsprozess für DiGAs aus den Vorarbeiten des AIHTAs 2020 wurden die Anforderungen des Allgemeinen Sozialversicherungsgesetzes (ASVG) und die Voraussetzungen für eine Kompatibilität der DiGAs mit der elektronischen Gesundheitsakte (ELGA) bisher nicht berücksichtigt; da beide Thematiken für den Umgang mit DiGAs jedoch wesentlich sind, werden sie im vorliegenden Bericht entsprechend miteinbezogen.

Dieser Prozess soll nun anhand einer digitalen Gesundheitsanwendung, einem sogenannten „Symptom-Checker“ pilotiert werden. Hier eine kurze Erläuterung, welche Ziele „Symptom-Checker“ verfolgen und wie sich diese Anwendungen von anderen, ähnlichen Systemen unterscheiden.

1.3 Symptom-Checker

**Symptom-Checker:
Unterstützung bei
Diagnosefindung,
Triage-Entscheidung**

Nutzer*innen recherchieren zunehmend im Internet nach möglichen Ursachen für bestehende Symptome. Neben Suchen in gängigen Suchmaschinen wie Google oder Bing gibt es ein wachsendes Angebot an diagnostischen Gesundheitsanwendungen, die bei der Diagnosefindung helfen sollen [14]. Diese auf Algorithmen basierenden diagnostischen Gesundheitsanwendungen werden als „Symptom-Checker“ bezeichnet. Sie stellen den Nutzer*innen nach Erfassen der Symptome mögliche ursächliche Diagnosen vor, und/oder geben Empfehlungen für ein entsprechendes Handeln (Selbstversorgung, Allgemeinmedizin oder Notfalls-Ambulanz) [9].

**CDSS:
Verwendung
ausschließlich durch
Gesundheitspersonal**

Von Symptom-Checkern abzugrenzen sind diagnostische Werkzeuge, die sich ausschließlich an das Gesundheitspersonal wenden, sog. „computerised diagnostic decision support programmes“ (CDDS) [15] bzw. „computerised clinical decision support systems“ (CDSS) Manche der verfügbaren Symptom-Checker waren ursprünglich als Unterstützung für Gesundheitspersonal (als CDSS) konzipiert, entwickelten sich mit der Zeit jedoch in Richtung einer anderen Anwendergruppe: den Nutzer*innen und Patient*innen [16-18]. Diagnostische Werkzeuge welche computergestützte Bildauswertungen durchführen, sind ebenfalls von Symptom-Checkern abzugrenzen. Ein Beispiel hierfür wäre eine DiGA, welche über das Smartphone Muttermale fotografiert und nach Hautkrebsrisiko analysiert.

**1999:
erster Symptom-Checker
veröffentlicht**

Der wohl erste entwickelte Symptom-Checker war „Isabel“, welcher mit Grundfunktionen bereits im Jahr 1999 veröffentlicht wurde [19]. Seither nahm sowohl die Zahl als auch die Nutzung der verfügbaren Symptom-Checker zu. Die Funktionsweise der Algorithmen verschiedener Symptom-Checker unterscheiden sich, in Verwendung kommen etwa „branching logic“, „bayesian inference“, „machine learning“ und andere Methoden [14, 20].

**webbasiert und/oder
Smartphone-App:
Chatbot-Funktion**

Meist sind Symptom-Checker über Websites aufrufbar, einige der Symptom-Checker sind zusätzlich als App für Smartphone oder Tablet verfügbar. Symptom-Checker sind häufig dialog-basiert (mittels Chatbot) aufgebaut, das heißt in Abhängigkeit der angegebenen Symptome werden Folgefragen gestellt [14].

**vergleichbare Ziele:
telefonische
Gesundheitsnummern**

**in Ö seit Nov. 2019:
„1450 – die
Gesundheitsnummer“**

Ein häufig genannter Vergleich für Symptom-Checker sind telefonische Gesundheitsberatungen, welche in westlichen Ländern zunehmend im Einsatz sind [14]. In Österreich ist das die „Gesundheitsnummer 1450“, welche seit November 2019 österreichweit verfügbar ist. Sie wurde nach dem Vorbild ähnlicher telefonischer Triage-Dienste in Großbritannien, Dänemark und der Schweiz entwickelt. Die Gesundheitsnummer 1450 ist rund um die Uhr verfügbar und bietet den Anrufer*innen basierend auf der geschilderten Symptomatik eine Behandlungsempfehlung sowie eine Einstufung der Dringlichkeit (Selbstbehandlung, Hausarzt*in, fachärztliche Behandlung oder Notfalls-Ambulanz). Die Einschätzung erfolgt durch speziell medizinisch geschultes diplomiertes Krankenpflegepersonal. Die Gesundheitsnummer 1450 wird vom Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz, den Bundesländern sowie den Sozialversicherungsträgern betrieben [21].

2 Methodisches Vorgehen

2.1 Projektziel und Forschungsfragen

Aufbauend auf einem bereits erarbeiteten groben Evaluationsmodell ist das Projektziel dieser Arbeit, den Bewertungsprozess für DiGAs in Österreich weiter zu konzipieren und diesen anhand eines Beispiels zu pilotieren.

Folgende Forschungsfragen sollen beantwortet werden:

- **FF1:** Wie kann die Vielzahl der verfügbaren digitalen Gesundheitsanwendungen unter Berücksichtigung des bereits vorliegenden Bewertungsframeworks [6] vor einer Evaluierung zusätzlich kategorisiert und priorisiert werden?
- **FF2:** Welche technologie-spezifischen Anforderungen müssen DiGA erfüllen, bevor sie einer umfassenden Evaluation zum Nutzen unterzogen werden?
- **FF3:** Welche Evidenz zum (medizinischen oder organisatorischen) Nutzen liegt für die Gruppe der „Symptom Checker“ vor und wie ist diese im Kontext der Kategorisierung und der Anforderungen zu bewerten?
- **FF4:** Welche weiteren Prozessschritte sind nach Adaption des Bewertungsframeworks in einer Feinkonzeption zu integrieren?

Rationale FF1: Die Kategorisierung und Priorisierung von DiGAs wird als sinnvoll erachtet, um aus der Vielzahl der verfügbaren DiGAs eine strukturierte Auswahl für Health Technology Assessments auszuwählen.

**Projektziel:
verfeinerte
Prozesskonzeption
für DiGAs**

4 Forschungsfragen

2.2 Übersicht zu digitalen Gesundheitsanwendungen, Priorisierung und technologische Anforderungen

Zur Beantwortung der ersten und zweiten Forschungsfrage wurde eine Handsuche, Expert*innenkonsultation und Datenextraktion priorisierter DiGAs durchgeführt:

Erstellung eines Meta-Verzeichnisses und Kategorisierung:

Zur Identifikation derzeitig verfügbarer DiGAs wurde eine Handsuche in sieben DiGA Verzeichnissen aus vier Ländern (Belgien, Deutschland, Vereinigtes Königreich, Österreich) durchgeführt. Die Handsuche zur Identifikation von DiGAs wurde durch drei Expert*innenkonsultationen ergänzt. Für DiGAs, die somit vor 30.07.2021 identifiziert wurden, erfolgte eine Datenextraktion der Sprache, Anwender (Nutzer*innen und/oder Gesundheitspersonal), Zielgruppe (Gesundheitszustand/Vulnerabilität der Nutzer*innen), Funktion der DiGA und Indikation. Zielsetzung des Meta-Verzeichnisses war es, DiGAs die in Europa zur Verfügung stehen zu sammeln um einerseits einen Überblick über die Verteilung in verschiedene Kategorien zu geben, und andererseits Aussagen zu treffen, welche der verfügbaren DiGAs für Österreich sinnvoll sein könnten.

**Handsuche, Konsultation
mit Expert*innen,
Datenextraktion**

**Erstellung
Meta-Verzeichnis
von DiGAs**

Folgende Institutionen wurden bezüglich DiGA am österreichischen Markt befragt:

1. 18.05.2021: AIT Austrian Institute of Technology GmbH [22]
2. 23.06.2021: DSVV Dachverband der Sozialversicherungsträger [23]
3. 28.06.2021: INiTS Universitäres Gründerservice Wien GmbH (Health Hub Vienna) [24]

Erarbeitung von Filter- und Priorisierungskriterien:

Kategoriensystem & Priorisierungskriterien

Aufbauend auf einem vom Fachgebiet Management im Gesundheitswesen (Technische Universität Berlin) erstellten Kategoriensystem wurde jede DiGA nach Zielgruppe, Funktion der DiGA und Indikation kategorisiert. Zusätzlich wurden zwei Expert*innenkonsultationen durchgeführt, um für Österreich relevante Priorisierungskriterien zu erstellen und zu pilotieren.

Folgende Expert*innenkonsultationen wurden bezüglich Priorisierung durchgeführt:

1. 23.06.2021: DSVV Dachverband der Sozialversicherungsträger [23]
2. 19.07.2021: DSVV Dachverband der Sozialversicherungsträger [23]

Überprüfung der technologie-spezifischen Anforderungen:

DiGA-spezifische (technische) Anforderungen

Die aus der Priorisierung gefilterten DiGAs wurden folgendermaßen beschrieben: Name und Beschreibung der DiGA, Verfügbarkeit (webbasiert und/oder App Stores), Verzeichnis, Anwender*innen, Zielgruppe, Funktionskategorie sowie Indikation. Anschließend erfolgte eine Übersicht zu folgenden technologie-spezifischen Anforderungen:

- CE-Kennzeichnung,
- Risikoklasse (nach Medizinproduktverordnung 2017/745 MDR),
- Angaben zum Datenschutz,
- Angaben zu verwendeten Algorithmen,
- Evidenzstufe (nach NICE Evidence Standards Framework [11]),
- Interoperabilität (verwendete technische Standards),
- ELGA-Kompatibilität

Diese technischen Charakteristika sollten idealerweise geprüft werden, bevor die DiGAs einer umfassenden Evaluation zum Nutzen unterzogen werden.

Zwei Expert*innenkonsultationen wurden zusätzlich durchgeführt, um den österreichischen Kontext hinsichtlich einer potentiellen zukünftigen Interoperabilität neuer DiGAs zu reflektieren:

1. 17.05.2021: ITSV IT-Services der Sozialversicherung GmbH [25]
2. 17.09.2021: ELGA Elektronische Gesundheitsakte GmbH [26]

2.3 Systematische Übersicht zu „Symptom-Checkern“

Zur Beantwortung der dritten Forschungsfrage wurde eine systematische Literatursuche durchgeführt und eine systematische Übersicht zu Ergebnissen zu DiGAs aus der Gruppe der Symptom-Checker erstellt. Es wurde bereits in der orientierenden Handsuche eine systematische Übersichtsarbeit des National Institute for Health Research identifiziert [9]. Diese ist 2019 publiziert (Literatursuche bis April 2018): diese wurde mit rezenter Literatur ergänzt.

Update eines SR (2019) über Symptom-Checker

2.3.1 Literatursuche

Die systematische Literatursuche dient einer Aktualisierung dieser systematischen Übersichtsarbeit. Die Literatursuche wurde am 20.05.2021 in den folgenden vier Datenbanken durchgeführt:

Literatursuche in 4 Datenbanken

- Cochrane Datenbank
- Embase Datenbank
- Medline Datenbank via Ovid
- Centre for Reviews and Dissemination (CRD) Datenbank, resp. INAHTA Datenbank

Die verwendeten Suchstrategien im Detail finden sich im Appendix Kapitel „Suchstrategien“.

Ein- und Ausschlusskriterien

Tabelle 2-1: Einschlusskriterien für die systematische Literatursuche

Population	<ul style="list-style-type: none"> ■ Gesundheitspersonal ■ Symptomatische Individuen („potentielle Patient*innen“) <p>Rationale: Beide Anwendungsszenarien wurden gewählt, da manche der verfügbaren Symptom-Checker ursprünglich als Unterstützung für Gesundheitspersonal konzipiert waren, sich mit der Zeit jedoch in Richtung Anwendergruppe Patient*innen entwickelten [16-18].</p>
Intervention	<p>Symptom-Checker (webbasiert und/oder Smartphone-App):</p> <ul style="list-style-type: none"> ■ Als Hilfestellung zur Steuerung von Patient*innenflüssen (Anwender: Gesundheitspersonal) ■ Als Hilfestellung symptomatischer Individuen (Anwender: Versicherte/„potentielle Patient*in“)
Comparators	<ul style="list-style-type: none"> ■ herkömmliche Versorgung/Standard of Care (SoC) zur Steuerung von Patient*innenflüssen (z. B. Telefon-Triage) ■ herkömmliche Versorgung/Standard of Care (SoC) (mit oder ohne Arztbesuch) symptomatischer Individuen ■ (Studien mit anderen Komparatoren (oder ohne Komparator), wenn sie die Forschungsfrage betreffen)
Outcomes (Endpunkte)	<p>Klinische Wirksamkeit</p> <ul style="list-style-type: none"> ■ Krankheitsdauer, Schwere der Erkrankung <p>Sicherheit</p> <ul style="list-style-type: none"> ■ Schäden durch falsch-positive/falsch negative Diagnosevorschläge oder falsche Triage-Empfehlungen <p>Genauigkeit der Diagnosevorschläge</p> <ul style="list-style-type: none"> ■ Sensitivität ■ Spezifität ■ Positiv prädiktiver Wert ■ Negativ prädiktiver Wert ■ Übereinstimmung mit klinischer Diagnose ■ Diagnosezeitpunkt

Outcomes (Endpunkte) (Fortsetzung)	<p>Genauigkeit der Triage-Empfehlungen</p> <ul style="list-style-type: none"> ■ Sensitivität ■ Spezifität ■ Positiv prädiktiver Wert ■ Negativ prädiktiver Wert ■ Sicherheit/Angemessenheit der Triage-Empfehlungen <p>Systemisch/organisatorische Endpunkte</p> <ul style="list-style-type: none"> ■ Anzahl Arzt-/Ambulanzbesuche <p>Patient*innenrelevante Endpunkte</p> <ul style="list-style-type: none"> ■ Lebensqualität, Zufriedenheit der Patient*innen ■ Bedienbarkeit der digitalen Gesundheitsanwendung ■ Haltung der Patient*innen gegenüber Symptom-Checkern <p>Rationale: Relevante Endpunkte wurden auf Basis einer systematischen Übersichtsarbeit sowie nach den „MMA Module for HTA“ und „Digi-HTA“ Bewertungsinstrumenten gewählt [9, 12, 13].</p>
Studiendesign	<ul style="list-style-type: none"> ■ Randomisierte kontrollierte Studien (RCTs) ■ Nicht-randomisierte kontrollierte Studien (NRCTs) ■ Beobachtungsstudien, Registerstudien, „Real-World-Evidence“ ■ Übersichtsarbeiten ■ Evaluierungsberichte
Publikationszeitraum	<p>Publikationen seit April 2018</p> <p>Rationale: Suchzeitraum der Systematischen Übersichtsarbeit des NIHR [9] bis April 2018</p>
Sprache	Deutsch, Englisch

Abkürzungen: HTA: Health Technology Assessment; MMA: digitale Gesundheitsanwendung („Mobile Medical Application“); NIHR: National Institute of Health Research; NRCT: nicht-randomisiert kontrollierte Studie („non-randomized controlled trial“); RCT: randomisiert kontrollierte Studie („randomized controlled trial“); SoC: herkömmliche Versorgung („Standard of Care“)

Ausschlusskriterien:
keine diagnose-spezifischen Symptom-Checker

Ausschlusskriterien:

- Diagnose-spezifische Symptom-Checker (z. B. für COVID-19, Migräne oder Dorsalgie)
- Qualitative Studien zu sekundären Daten (z. B. auf Basis von App-Bewertungen in den App Stores)
- Qualitative Studien zur Gesundheitspersonal-Perspektive („physician perspective surveys“)
- Studien aus Ländern mit für Österreich nicht vergleichbaren Gesundheitssystemen

Literaturauswahl

1.300 Referenzen:
SR mit 14 Studien
„updated“

Insgesamt konnten 1.300 Quellen durch die systematische Literatursuche identifiziert werden, darunter eine durch die erweiterte Handsuche. Nach der Deduplizierung resultierten beide Suchen in 1.009 Treffer. Die Literatur wurde unter Berücksichtigung der vordefinierten Ein- und Ausschlusskriterien ausgewählt. Nach dem Abstract-Screening dieser Treffer wurden 122 Volltexte für die weitere Durchsicht ausgewählt. Basierend auf der Volltextdurchsicht wurden in Ergänzung zur systematischen Übersichtsarbeit des National Institute for Health Research [5] 14 weitere Publikationen (acht Beobachtungsstudien sowie sechs qualitative Studien für die Patient*innensicht) eingeschlossen. Das Abstract- und Volltext-Screening wurde von zwei Personen durchgeführt (RJ, GG). Differenzen wurden durch Diskussion und Konsens gelöst. Der Auswahlprozess ist in Abbildung 2-1 dargestellt.

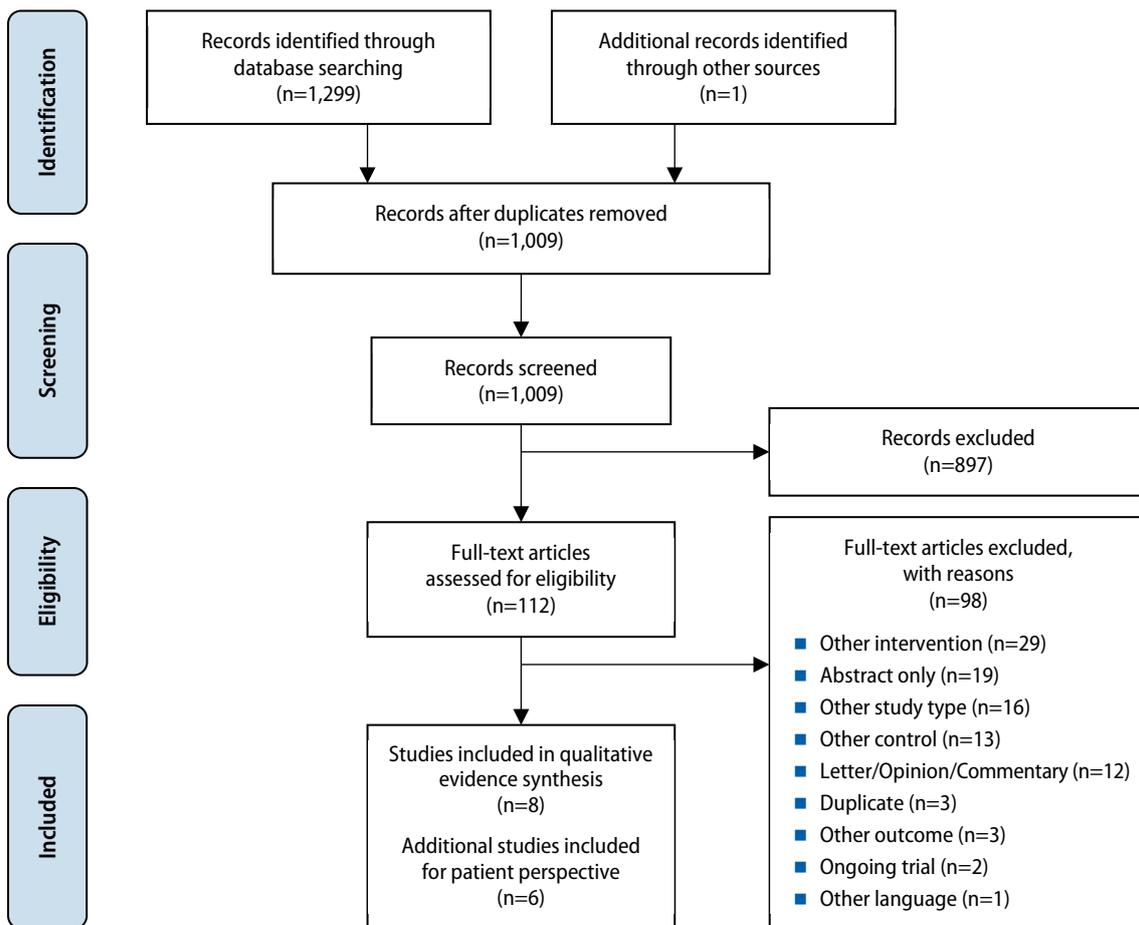


Abbildung 2-1: Darstellung des Auswahlprozesses (PRISMA Flow Diagram)

2.3.2 Analyse und Qualitätsbewertung

Die systematische Übersichtsarbeit des National Institute for Health Research [9] wurde mit dem AMSTAR-II Assessment Tool [27] bewertet (vgl. Appendix „Qualitätsbewertung des NIHR Berichts mit dem AMSTAR-II Assessment Tool“).

Qualitätsbewertung & RoB-Bewertung

Die acht zusätzlich identifizierten Beobachtungsstudien für die Evidenzsynthese wurden mit dem QUADAS-2 Assessment Tool [28] für Risk-of-Bias von diagnostischen Studien bewertet (vgl. Appendix „Risk-of-Bias Assessment für diagnostische Studien mit dem QUADAS-II Assessment Tool“). Für die sechs qualitativen Studien für die Patient*innensicht wurde kein Risk-of-Bias Assessment durchgeführt.

Alle Analysen wurden von einem Wissenschaftler (RJ) durchgeführt und von einem zweiten Wissenschaftler (GG) kontrolliert.

2.3.3 Datenextraktion und Synthese

Datenextraktion relevanter Endpunkte

Im Anschluss an die Literaturbeschaffung erfolgte die Datenextraktion aus den eingeschlossenen Studien. Diese beinhaltete neben der Studien- und Patient*innen-Charakteristika die Endpunkte klinische Wirksamkeit, Sicherheit, Genauigkeit der Diagnosevorschläge, Genauigkeit der Triage-Empfehlungen, systemisch/organisatorische Endpunkte sowie patient*innenrelevante Endpunkte. Die Datenextraktion und narrative Beschreibung der Evidenz wurden von einem Wissenschaftler (RJ) durchgeführt und von einem zweiten Wissenschaftler (GG) kontrolliert.

2.3.4 Endpunkte der eingeschlossenen Studien

Es wurden folgende Endpunkte als relevant erachtet:

Wirksamkeit

Krankheitsdauer und Schwere der Erkrankung

Für die klinische Wirksamkeit der Symptom-Checker wurde Evidenz zu Krankheitsdauer und Schwere der Erkrankung bei Verwendung der Intervention definiert.

Sicherheit

Schäden durch falsch-positiv/ falsch-negativ Diagnosevorschläge oder falsche Triage- Empfehlungen

Für die Sicherheit der Symptom-Checker wurde Evidenz zu Schäden durch falsch-positiv/falsch-negativ Diagnosevorschläge oder falsche Triage-Empfehlungen definiert: etwaige Schäden durch unnötige Untersuchungen oder Folgeeingriffe bei falsch-positiven Diagnosevorschlägen oder übervorsichtigen Triage-Empfehlungen, sowie etwaige Schäden durch Nicht-Behandlung bei falsch-negativen Diagnosevorschlägen oder vermessenen Triage-Empfehlungen.

Genauigkeit der Diagnosevorschläge

diagnostische Genauigkeit: Sensitivität, Spezifität, positiv/negativ prädiktiver Wert

Für die diagnostische Genauigkeit wurden die Endpunkte Sensitivität, Spezifität, positiv prädiktiver Wert und negativ prädiktiver Wert definiert.

Genauigkeit der Triage-Empfehlungen

Genauigkeit Triage: Sensitivität, Spezifität, positiv/negativ prädiktiver Wert

Für die Genauigkeit der Triage-Empfehlungen wurden die Endpunkte Sensitivität, Spezifität, positiv prädiktiver Wert und negativ prädiktiver Wert definiert. Außerdem zählen zur Genauigkeit der Triage-Empfehlungen die Sicherheit/Angemessenheit der Triage-Empfehlungen: ob die genannte Triage-Empfehlung für den klinischen Fall als angemessen angesehen wird.

Systemisch/organisatorische Endpunkte

Anzahl der Arzt- /Ambulanzbesuche

Für diesen Endpunkt wurde Evidenz zur Auswirkung der Symptom-Checker auf die Anzahl der Arzt-/Ambulanzbesuche definiert.

Patient*innenrelevante Endpunkte

Für die Patient*innensicht wurden die Endpunkte Lebensqualität/Zufriedenheit der Patient*innen definiert. Außerdem Evidenz zur Anwenderfreundlichkeit der Symptom-Checker sowie zur Haltung der Patient*innen gegenüber Symptom-Checkern.

**Lebensqualität/
Zufriedenheit
Anwender-freundlichkeit**

2.4 Pilotevaluierung des Prozesses anhand einer ausgewählten DiGA („Symptoma“)

Zur Beantwortung der dritten Forschungsfrage erfolgte exemplarisch eine Pilotevaluierung des Prozesses. Hierfür wurde vonseiten des Auftraggebers die DiGA „Symptoma“ aus der Gruppe der Symptom-Checker [5] ausgewählt. Dazu wurde ein Abgleich mit den hier vorgeschlagenen Prozessschritten zur ASVG-Relevanz, sowie technologischer Anforderungen und den im Teil 1 [6] erarbeiteten Evidenz- und Evaluationsanforderungen des Evaluationsmodells (Evidenzanforderungen: NICE Evidence Standards Framework [9]; Evaluationsanforderungen: HTA Module for Mobile Medical Applications [10], DiGi-HTA Framework [11]) durchgeführt.

**Pilotevaluierung
des in Teil 1 erarbeiteten
Prozesses**

In der nachfolgenden Tabelle 2-2 ist eine Einteilung der DiGAs in verschiedene Risikoklassen nach EU-Verordnung 2017/745 über Medizinprodukte gegeben [10]. Diese Einteilung ist eine Voraussetzung für die regulatorische Bewertung einer DiGA.

Tabelle 2-2: Risikoklassen der digitalen Gesundheitsanwendungen [10]

Risikoklassen nach EU-Verordnung 2017/745 über Medizinprodukte (Regel 11)
<p>Risikoklasse 1 Jede Software, die nicht in Risikoklasse 2a oder höher eingeteilt wird</p>
<p>Risikoklasse 2a Software welche Information zur Entscheidung einer Diagnose oder therapeutischen Zwecken zur Verfügung stellt, ohne jedoch dabei ernstere Gesundheitsschäden verursachen zu können (siehe nachfolgende Risikoklassen). Software die physiologische Messwerte aufzeichnet, solange es sich dabei nicht um Vitalparameter handelt</p>
<p>Risikoklasse 2b Software welche Information zur Entscheidung einer Diagnose oder therapeutischen Zwecken zur Verfügung stellt, bei der ernstere Gesundheitsschäden oder die Notwendigkeit einer chirurgischen Intervention entstehen kann. Software die Vitalparameter aufzeichnet (Messwerte, bei denen eine Änderung eine unmittelbare Gefahr für die Patient*innen darstellen kann)</p>
<p>Risikoklasse 3 Software welche Information zur Entscheidung einer Diagnose oder therapeutischen Zwecken zur Verfügung stellt, bei der diese Entscheidung ein Todesereignis oder irreversible Schäden der Gesundheit verursachen kann.</p>

2.5 Feinkonzeption des Prozesses

Prozess in Feinkonzeption adaptiert

Zur Beantwortung der vierten Forschungsfrage wurden die gewonnenen Erkenntnisse aus den Expert*innenkonsultationen, sowie aus der Identifikation, Kategorisierung und Priorisierung von DiGAs, den technologiespezifischen Anforderungen sowie der Pilotevaluierung des Prozesses anhand einer DiGA aus der Gruppe der Symptom-Checker von drei Personen (CW, RJ, GG) analysiert. Anschließend wurde der in Teil 1 (2020 [6]) skizzierte Bewertungsprozess für die Bewertung von DiGAs in Österreich mit weiteren Prozessschritten adaptiert.

2.6 Qualitätssicherung

Begutachtung durch 1 interne und 2 externe Reviewer*innen

Dieser Bericht wurde durch eine*n interne und zwei externe Reviewer*innen begutachtet (siehe „Projektbeteiligung“ auf Seite 2). Die externen Reviewer*innen wurden vor allem gebeten, die folgenden Qualitätskriterien zu bewerten:

- Fachliche Korrektheit: Ist der Bericht fachlich korrekt (verwendete Evidenz und Information)?
- Berücksichtigt der Bericht die neuesten Erkenntnisse in diesem Forschungsbereich?
- Angemessenheit und Transparenz der Methode: ist die gewählte Methode zur Beantwortung der Forschungsfrage geeignet und werden die Methoden transparent angewendet?
- Logische Struktur und Konsistenz des Berichts: Ist die Struktur des Berichts konsistent und verständlich?
- Formale Merkmale: Erfüllt der Bericht formale Kriterien des wissenschaftlichen Schreibens (z. B. korrekte Zitate)?

Das AIHTA betrachtet den externen Review-Prozess durch wissenschaftliche Expert*innen als eine Methode der Qualitätssicherung der wissenschaftlichen Arbeit. Die Entscheidung über die Berücksichtigung der Kommentare der Reviewer*innen in der Endfassung des Berichts liegt bei den Autor*innen des Berichts.

3 Ergebnisse

3.1 Übersicht zu digitalen Gesundheitsanwendungen und deren Kategorisierung wie Priorisierung

In folgendem Kapitel sollen zunächst Ansätze für eine Kategorisierung beschrieben werden. Diese basieren auf

- Vorarbeiten zu Assessment Frameworks für DiGAs des AIHTAs 2020 [6] und
- „iDiGA-Projekt“ des Fachgebiets (FG) Management im Gesundheitswesen (Technische Universität Berlin)¹ [30].

Darauf aufbauend wurde ein Meta-Register von ausgewählten europäischen DiGA-Verzeichnissen erstellt sowie Priorisierungskriterien zur Auswahl relevanter DiGAs (nach dem allgemeinen Sozialversicherungsgesetz, ASVG) [31] für eine Evaluation in Österreich erarbeitet.

3.1.1 Kategorisierung von DiGAs

In den Vorarbeiten zu Assessment Frameworks für DiGAs des AIHTAs 2020 [6] wurde die Taxonomie der World Health Organization (WHO) zur Kategorisierung von digitalen und mobilen Technologien beschrieben („Classification of Digital Health Interventions“) [32]. Diese Taxonomie unterscheidet grob vier Bereiche, je nachdem ob die Technologie von den Nutzer*innen oder durch die behandelnden Arzt*innen oder Therapeut*innen verwendet wird. Technologien für Datendienste, sowie Technologien die dem Management im Gesundheitswesen auf Systemebene dienen, bilden jeweils eigene Kategorien.

Nach dieser Taxonomie der WHO sind die meisten DiGAs der ersten Kategorie zuzuordnen, da sie von den Nutzer*innen selbst eingesetzt werden. Von diesen DiGAs abzugrenzen sind Anwendungen auf Gesundheitsanbieter-Ebene oder Anwendungen auf der Gesundheitssystem-Ebene.

In den Vorarbeiten des AIHTAs 2020 [6] wurden ausgewählte DiGAs grob nach jeweiligem Output (informativ, diagnostisch und/oder therapeutisch) eingeteilt. Eine weiter granulierte Kategorisierung nach Funktion der DiGAs ist im Bericht des Forschungsprojektes „iDiGA“, geleitet vom Fachgebiet Management im Gesundheitswesen der Technischen Universität Berlin [30] beschrieben. Auf Basis dieses Ansatzes können DiGAs anhand von vier Kriterien kategorisiert werden:

Anwendungsbereich: Der Anwendungsbereich ergibt sich im Wesentlichen aus dem 3-stelligen ICD-10 Code der internationalen Klassifikation der Krankheiten.

Zielgruppe: Die Zielgruppe ist nach der „Vulnerabilität“ der Patient*innen in fünf Ausprägungen untergliedert. Demnach kann sich eine DiGA an Gesunde ohne bekannte Risikofaktoren (0), Gesunde mit Risikofaktoren (1),

Kategoriensystem:
basierend auf
Projektbericht Teil 1
und TU Berlin

WHO Taxonomie:
DiGAs von Nutzer*innen
eingesetzt

AIHTA Projektbericht Teil 1:
Output = informativ,
diagnostisch und/oder
therapeutisch

iDiGA-Projekt Kategorien:
Anwendungsbereich,
Zielgruppe, Funktion und
Anwender*innen

¹ Dieser Kategorisierungsansatz wurde wiederum durch Vorarbeiten der Bertelsmann Stiftung [29] und dem NICE Evidence Standards Framework [11] informiert.

nicht lebensbedrohlich akut Erkrankte (2), chronisch Erkrankte mit stabilem Gesundheitszustand (3) oder Patient*innen mit instabilem Gesundheitszustand richten.

0. **Gesund ohne bekannte Risikofaktoren:** Es handelt sich um Personen ohne bekannte Risikofaktoren, die nicht durch eine Diagnose als „vorerkrankt“ gelten. Eine DiGA, welche ihre Zielgruppe nicht klar bestimmt (z. B. eine DiGA, die „alle Menschen“ mit oder ohne medizinisch relevante Symptome unterstützen soll) ist im Zweifelsfall dieser Zielgruppe zuzuordnen.
1. **Gesunde mit Risikofaktoren:** Für eine Zuordnung in diese Zielgruppe müssen bereits bekannte Risikofaktoren (z. B. Vorerkrankungen) vorliegen.
2. **Akut erkrankt, nicht lebensbedrohend:** Diese Zielgruppe beinhaltet all jene Patient*innen, die zwar akut erkrankt sind, für welche jedoch keine Notfallversorgung für die Genesung erforderlich ist.
3. **Chronisch erkrankt, stabiler Gesundheitszustand:** Diese Zielgruppe umfasst Personen mit einer diagnostizierten chronischen Erkrankung, die jedoch in keinem lebensbedrohlichen Zustand (stabiler Gesundheitszustand) sind.
4. **Hochvulnerabel, instabiler Gesundheitszustand:** Diese Zielgruppe umfasst hochvulnerable Patient*innen (z. B. schwanger, abweichende/s Wahrnehmung, Denken, Fühlen oder Selbstbild, Pflegebedürftige).

Funktion: Die Funktion einer DiGA kann in drei Kernfunktionen gegliedert werden: Erkennung, Überwachung und Behandlung². Anhand einer genaueren Funktionalität können diese noch weiter untergliedert werden (vgl. Tabelle 3-1). Die granulare Untergliederung richtet sich dabei an der europäischen Medizinprodukteverordnung (Art. 2 Nr. 1): „(2) Verhütung, Vorhersage oder Prognose und (3) Untersuchung eines physiologischen oder pathologischen Vorgangs oder Zustands“. Für eine breite Definition einer DiGA wurde (0) Dokumentation und (1) Information als weitere Gliederungen hinzugefügt, wiewohl diese über die Zweckbestimmung gemäß MDR hinausgehen.

Anwender*innen: DiGAs können entweder nur von Patient*innen (bzw. auch Angehörigen) oder von Anwender*innen und Leistungserbringer genutzt werden. Die ausschließliche Verwendung einer DiGA von Leistungserbringern stellt einen weiteren Anwendungsfall dar.

Eine Übersicht über diesen Kategorisierungsansatz anhand der oben beschriebenen vier Kriterien – mit den jeweils möglichen Ausprägungen – ist in der nachfolgenden Tabelle 3-1 dargestellt.

² Diese orientieren sich an den im deutschen Sozialgesetzbuch geregelten zulässigen Zweckbestimmungen (§ 33a Abs. 1 SGB V).

Tabelle 3-1: Kategorisierungsansatz von DiGAs übernommen von Vorarbeiten des iDIGA-Projektes [30]

Anwendungsbereich (nach ICD 10) – Beispiele	Zielgruppe	Funktion der DiGA*		AnwenderIn
...	0 Gesund ohne bekannte Risikofaktoren	0 Dokumentation (NA)		1 Nur Patient*innen bzw. Angehörige
Exx Endokrine, Ernährungs- und Stoffwechsel- Krankheiten	1 Gesund mit Risikofaktoren	1 Information (NA)		2 Patient*innen & Leistungserbringer
Fxx Psychische und Verhaltensstörungen	2 akut erkrankt, nicht lebensbedrohend	2 Verhütung, Vorhersage, Prognose (NA)		3 Nur Leistungserbringer
Hxx Krankheiten des Auges und des Ohres	3 chronisch erkrankt, stabiler Gesundheitszustand	3 Untersuchung eines physiologischen oder pathologischen Vorgangs oder Zustands (NA)		
Ixx Krankheiten des Kreislaufsystems	4 hochvulnerabel/instabiler Gesundheitszustand	Erkennung (Diagnose)	4 Diagnosestellung	
Oxx Schwangerschaft, Geburt und Wochenbett		Überwachung (Monitoring)	5 Einfaches Monitoring	
			6 Komplexes Monitoring	
...		Behandlung (Therapie)	7 Indirekte Intervention: Selbst-Management	
			8 Direkte Intervention: Veränderung von Gesundheitszuständen	

Quelle: [30]

Legende: DiGA: Digitale Gesundheitsanwendung; ICD-10: International Classification of Diseases, 10. Ausgabe; NA: Nicht zutreffend (not applicable); TU Berlin: Technische Universität Berlin

*Funktion der DiGA: Diagnosestellung: DiGA, die dazu dienen, bestimmte Zustände zu erkennen oder Leistungserbringer (LE)/andere Akteure bei der Diagnosestellung zu unterstützen (und LE nicht ersetzen). Einfaches Monitoring: DiGA, die kontinuierlich Körperfunktionen/Zustände messen, welche nicht direkt überlebenswichtig oder schwerwiegende therapeutische Konsequenzen zur Folge haben können. Die Daten werden in der Regel nicht automatisch an den LE/andere Akteure übermittelt. Komplexes Monitoring: DiGA, die kontinuierlich Körperfunktionen/Zustände messen, die überlebenswichtig sind oder schwerwiegende therapeutische Konsequenzen zur Folge haben können. Die Daten werden ggf. automatisch an LE/andere Akteure übermittelt. Indirekte Intervention (Selbst-Management): DiGA, die Anwender*in mit einer diagnostizierten Krankheit zu einer besseren Selbststeuerung anleiten und diese aktiv zu einer eigenständigen Problembewältigung befähigen, um Therapieziele zu erreichen. Hierunter fallen insbesondere solche DiGA, die die bestehende Versorgung ergänzen. Direkte Intervention (Veränderung von Gesundheitszuständen): DiGA, die direkt auf eine Veränderung von Gesundheitszuständen abzielen oder therapeutische Entscheidungen leiten oder einen diagnostizierten Zustand therapieren. Hierunter fallen insbesondere auch solche DiGA, die bestehende Versorgungskomponenten ersetzen.

Mögliche Komparatoren und erwartete Endpunkte einer Evaluierung

**für DiGAs geeignete
Komparatoren und
Endpunkte
(öffentlicher Diskurs)**

Neben einer Kategorisierung von DiGAs sollte die Überlegung passender Komparatoren sowie adäquater Endpunkte reflektiert werden, um in der Bewertung diejenigen DiGAs selektieren zu können, welche den größten gesundheitlichen Zusatznutzen zur Standardtherapie bringen.

Im **deutschen DiGA-Verzeichnis** wird es als akzeptabel erachtet, „keine Therapie“ als passenden Vergleich zu wählen, unabhängig davon, ob die DiGA dafür konzipiert ist, eine Versorgungslücke zu schließen. Als adäquate Endpunkte zur Evaluierung des medizinischen Nutzens werden patient*innen-relevante Endpunkte (wie Mortalität, Morbidität oder Lebensqualität) genannt. Surrogatendpunkte oder andere Endpunkte wie „Adhärenz“ können dann herangezogen werden, wenn DiGAs für die Struktur- und Verfahrensverbesserung konzipiert sind [4, 6].

Es ist darauf hinzuweisen, dass es international noch keinen Konsens darüber gibt, welche Komparatoren/Endpunkte innerhalb der Evaluation von DiGAs heranzuziehen sind. Komparatoren können „keine Therapie“ oder die „übliche Versorgung“ umfassen. Denkbar ist auch, dass eine bestehende Versorgungslücke eine Bedingung für den Vergleich „keine Therapie“ darstellt. Patient*innenrelevante Endpunkte (z. B. Verbesserung der Lebensqualität) sind unumstritten geeignete Endpunkte für die Evaluation von DiGAs [6]. Inwiefern Surrogatendpunkte oder andere Endpunkte (wie etwa die Adhärenz) einen Aufschluss über den (bspw. strukturverbessernden) Nutzen einer DiGA geben können bleibt Bestandteil eines wissenschaftlichen und öffentlichen Diskurses.

3.1.2 Erstellung eines Meta-Registers von DiGAs

**Meta-Verzeichnis
4 Länder:
AT, BE, DE, UK**

Um einen Überblick über die derzeitig verfügbaren DiGAs zu bekommen wurde zunächst ein Meta-Verzeichnis aus sieben Verzeichnissen/Auflistungen aus vier Ländern erstellt (siehe Appendix Kapitel „Meta-Verzeichnis von digitalen Gesundheitsanwendungen“). In Tabelle 3-2 befindet sich die Übersicht der ausgewählten Verzeichnisse sowie erste Refundierungsansätze der ausgewählten Länder. Eine detaillierte Beschreibung zu Refundierungsprozessen im Ländervergleich befindet sich im Projektbericht zu digitalen Gesundheitsanwendungen des AIHTAs 2020 [6].

Tabelle 3-2: Übersicht zu den im Meta-Verzeichnis inkludierten Verzeichnissen und Auflistung, angelehnt an den AIHTA Projektbericht 2020 [6]

Land	Deutschland	Großbritannien	Belgien	Österreich
Name des Verzeichnisses oder Auflistungen	DiGA Verzeichnis	NHS Apps Library	mHealth Belgium	IfGP + AIT, Health Hub Vienna Alumni; ScaleUp4Europe
Beschreibung	Positivliste erstatteter DiGAs	Listung aller (nachgewiesener sicherer) DiGAs ohne Rückschlüsse zu Erstattungsfähigkeit	Listung von DiGAs in einer sog. „Validierungspyramide“	Listung von DiGAs die in Österreich produziert oder erprobt wurden
Nationaler Erstattungsprozess vorhanden	Ja	Nein	geplant	Nein

Legende: DiGA-Verzeichnis: Verzeichnis für digitale Gesundheitsanwendungen; NHS Apps Library: National Health Service Apps Library; IfGP: Institut für Gesundheitsförderung und Prävention; AIT: Austrian Institute of Technology

Metaverzeichnis

Das Meta-Verzeichnis beinhaltet insgesamt 181 Einträge (nach De-Duplizierung: n=176). Davon erfüllten 132 Anwendungen die in diesem Bericht gewählte (bereits beschriebene) Definition einer DiGA.

**181 Einträge,
132 DiGAs**

Das Meta-Verzeichnis beinhaltet insgesamt 71 deutschsprachige DiGAs. Knapp über die Hälfte dieser DiGAs (n=39) richten sich an chronisch Erkrankte, knapp über ein Drittel dieser DiGAs (n=25) richten sich an Gesunde ohne bekannte Vorerkrankung. Wenige dieser DiGAs richten sich an Gesunde mit Risikofaktoren (n=5) oder akut erkrankte (n=2) und keine der identifizierten deutschsprachigen DiGAs richten sich an hochvulnerable Patient*innen (mit instabilem Gesundheitszustand).

**71 deutschsprachige
DiGAs – die Hälfte an
Chroniker;**

Die Verteilung der DiGAs nach ihrer Funktion zeigt, dass ein Großteil dieser im Bereich des Monitorings oder der Therapie ist (n=41). Weiters gibt es DiGAs mit informativen (n=22), diagnostischen (n=6) oder dokumentativen (n=2) Funktionen.

**n=41 – Therapie
n=22 – Information
n=6 – Diagnostik
n=2 – Dokumentation**

Die deutschsprachigen DiGAs des Meta-Verzeichnisses befinden sich im Anhang (siehe Tabelle A-6).

3.1.3 Priorisierungskriterien für die Evaluation von DiGAs in Österreich

Um aus der Vielzahl der verfügbaren DiGAs nach erfolgter Kategorisierung eine konkrete Anzahl an DiGAs für die Überprüfung der technologiespezifischen Anforderungen auszuwählen, sind neben dem oben beschriebenen Kategoriensystem weitere Kriterien für eine Priorisierung erforderlich. Die Erarbeitung dieser Priorisierungskriterien wird im nachfolgenden Kapitel beschrieben.

**Bedarf an
Priorisierungskriterien
(ASVG-Relevanz)**

In Anlehnung an das Allgemeine Sozialversicherungsgesetz (ASVG) [31] – und akkordiert mit entsprechenden Expert*innen – wurden Kriterien entwickelt, die den Anspruch erheben, aus der Fülle der derzeit verfügbaren DiGAs jene zu filtern, die unmittelbare ASVG-Relevanz aufweisen:

**deutschsprachig,
zur Behandlung von
chronisch Erkrankten**

- Sprache: Deutsch
- Zielgruppe: 3 – chronisch erkrankt, stabiler Gesundheitszustand
- Funktion der DiGA: Behandlung
(Therapie: 7 – indirekte Intervention und 8 – direkte Intervention)
- Anwender*in: (1) Nur Patient*innen bzw. Angehörige und (2) und Patient*innen und Angehörige, sowie Leistungserbringer.

Da die Auswahl der Priorisierungskriterien für DiGAs eng an die Bestimmungen des Sozialversicherungsrechts geknüpft sind, folgt nun ein Exkurs insbesondere zum Allgemeinen Sozialversicherungsgesetz (ASVG) in Bezug auf die Einordnung und Anerkennung von DiGAs im österreichischen Kontext.

Exkurs: Das Recht der sozialen Krankenversicherung

Welche Produkte und Dienstleistungen durch die österreichische Sozialversicherung als (Teil einer) Sachleistung gewährt bzw. erstattet werden, ist in zahlreichen und durchaus unterschiedlichen Regelungskomplexen normiert; zentral sind dabei die leistungsrechtlichen Bestimmungen des Allgemeinen Sozialversicherungsgesetzes (ASVG).

**Krankenbehandlung
ausreichend und
zweckmäßig, das Maß
des Notwendigen nicht
überschreitend**

Demgemäß muss eine Krankenbehandlung ausreichend und zweckmäßig sein, sie darf jedoch das Maß des Notwendigen nicht überschreiten. Durch die Krankenbehandlung sollen die Gesundheit, die Arbeitsfähigkeit und die Fähigkeit, für die lebenswichtigen persönlichen Bedürfnisse zu sorgen, nach Möglichkeit wiederhergestellt, gefestigt oder gebessert werden. (§ 133 Abs. 2 ASVG) [33]:

**DiGAs:
Heilbehelfe oder
„sonstige Heilmittel“**

Die DiGAs können – je nach Funktion und Ausgestaltung – als Heilbehelfe (§ 137 ASVG) oder als Heilmittel (§ 136 ASVG) qualifiziert werden. Bei letzterem kommt nur eine Einordnung abseits von Arzneimitteln (Medikamenten) als „sonstige Mittel, die zur Beseitigung oder Linderung der Krankheit oder zur Sicherung des Heilerfolges dienen“ wie z. B. Verbandstoffe in Betracht. (§ 136 Abs. 1 lit. b ASVG). Heilbehelfe wirken von außen auf die Besserung des Krankheitsbildes ein wie z. B. Schuheinlagen und Bandagen. Sie dienen der Heilung oder Linderung der Krankheit bzw. Verhütung von Verschlimmerungen der Krankheit.

**Pflichtleistung mit
stringenten Kriterien**

Im Bereich der Pflichtleistungen der Krankenversicherung, also dort wo auf die Krankenbehandlung ein (auch gerichtlich) durchsetzbarer Rechtsanspruch der Versicherten besteht, ist von einem sehr stringenten Regime auszugehen. Gerade die Kriterien der Zweckmäßigkeit (geeignet das Ziel der Krankenbehandlung zu erreichen), der Notwendigkeit (gibt es überhaupt einen entsprechenden Bedarf – dazu gehört auch das solidarisch zu garantierende Versorgungsniveau), Wirksamkeit (Nachweis durch Studien) und eben Kosteneffektivität sind einer Entscheidung zugrunde zu legen.

**Prävention,
Rehabilitation u. a. weit
weniger reguliert**

DiGAs, die ausschließlich Funktionen im Bereich der Gesundheitsförderung und Prävention beinhalten, stellen nach dem System des ASVG keine Pflichtleistungen der Krankenversicherung dar. Dieser weit weniger regulierte Bereich eignet sich auch besser dafür, Innovationen ausprobieren zu können. Ähnliches trifft z. B. auch auf die Rehabilitation, Maßnahmen zur Festigung der Gesundheit oder zur Stärkung der Gesundheitskompetenz usw. zu.

**keine Regelungen zur
Aufnahme in den
Leistungskatalog;
in Ö noch keine sozial-
versicherungs-rechtliche
DiGA-Beurteilung**

Allen Bereichen ist gemein, dass es grundsätzlich kein genau geregeltes Verfahren zur Aufnahme in das Leistungsangebot der Krankenversicherung gibt. Nur im enggesteckten Bereich der Heilmittel i. S. von Arzneien („Medikamente“) existieren für den Prüfprozess ausdrückliche und sehr fein zisierte gesetzliche Regelungen (inklusive einem „Rechtsschutz“ für die Anbieterfirmen) für die Aufnahme in den Erstattungskodex (EKO). Auch sonst gibt es grundsätzlich in Österreich derzeit noch keine abschließende rechtliche Beurteilung von DiGAs im System der sozialen Krankenversicherung.

**offene juristische
Fragen**

Die noch offenen und zu klärenden Fragen des ASVG in Bezug auf DiGAs sind in der Prozesskonzeption für die Bewertung von DiGA in Kapitel 3.5 näher beschrieben.

Anwendung der Priorisierungskriterien

**38 DiGAs erfüllen
Priorisierungskriterien
(Behandlung/Monitoring in
Chroniker-programmen)**

Nach Anwendung der in diesem Kapitel beschriebenen Kriterien (für Österreich relevante Kategorisierungs- und Priorisierungskriterien) waren für 38 von 181 identifizierten Einträge (nach De-Duplizierung: 176 Einträge) aus den oben genannten Verzeichnissen und Listungen die Priorisierungskriterien für eine weitere Überprüfung erfüllt.

132 Anwendungen erfüllten die in diesem Bericht gewählte (bereits beschriebene) Definition einer DiGA und 71 dieser DiGAs sind in Deutsch verfügbar. 39 der DiGAs wandten sich an chronisch erkrankte Patient*innen, wovon 38 DiGAs weitere Kriterien (Funktion: Behandlung oder Monitoring) erfüllten.

Eine Illustration dieser Anwendung auf die DiGAs (im Sinne eines Filters) ist in Abbildung 3-1 ersichtlich. Die Listung der 71 deutschsprachigen DiGAs (vor Anwendung der Priorisierungskriterien) findet sich im Meta-Verzeichnis im Appendix Tabelle A-6.

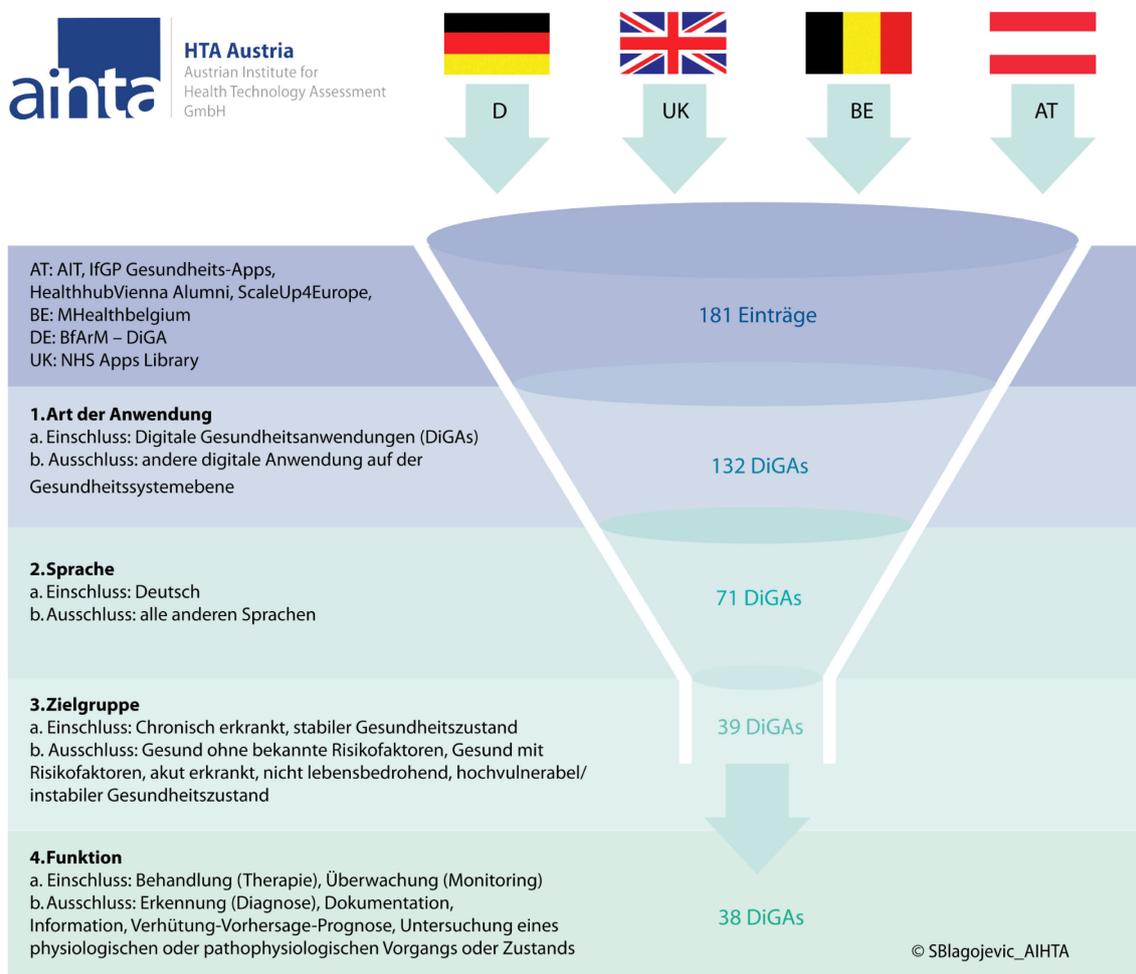


Abbildung 3-1: Priorisierung der für Österreich relevanten DiGAs

Notiz: Die Websites der jeweiligen Verzeichnisse befinden sich im Meta-Verzeichnis (siehe Tabelle A-6 im Anhang).

Abkürzungen: AIT – Austrian Institute for Technology; BfArM – Bundesinstitut für Arzneimittel und Medizinprodukte; IfGP – Institut für Gesundheitsförderung und Prävention.

3.2 Technologiespezifische Anforderungen an digitale Gesundheitsanwendungen

**technische Anforderungen,
Schnittstellen/
Interoperabilität/
Kompatibilität**

Unter Anwendung der in Kapitel 3.1.3 entwickelten Priorisierungskriterien (nach ASVG für Österreich unmittelbar relevante DiGAs) wurden insgesamt 38 DiGAs identifiziert. Um einen Überblick über die technologiespezifischen Anforderungen von DiGAs zu geben, folgt im nächsten Abschnitt zunächst ein Exkurs zu den technischen Voraussetzungen für eine Kompatibilität mit der österreichischen elektronischen Gesundheitsakte („ELGA“) und danach eine detailliertere Beschreibung der DiGAs und deren technischen Standards für Interoperabilität.

3.2.1 Technologische Anforderung für ELGA-Kompatibilität

**elektronische
Gesundheitsakte (ELGA)
noch nicht für DiGA
ausgerichtet**

Derzeit ist die österreichische elektronische Gesundheitsakte („ELGA“) aus zwei Gründen noch nicht für Schnittstellen zu DiGAs ausgerichtet:

- Die Zugriffsberechtigungen für das Einstellen von Dokumenten sind auf ELGA-Gesundheitsdiensteanbieter (ELGA GDA) gem. §2 Z10 GTeIG limitiert. DiGAs, die auf Basis einer Authentifizierung als Patient*in funktionieren, besitzen folglich keine Schreibberechtigung in ELGA.
- ELGA basiert auf dem Austausch von sogenannten CDA-Dokumenten. Es können somit nur Dokumente, nicht jedoch Einzelwerte, wie sie für Verlaufskurven etc. benötigt und von DiGAs vielfach erfasst werden, effizient in ELGA abgelegt werden.

Nachdem man sich aber dieser Problematik als limitierender Faktor für die Integration von DiGAs bewusst ist, wird bereits an Lösungen gearbeitet, entsprechende generische Schnittstellen für DiGAs anzubieten.

mHealth/DiGA Anbindung an ELGA:

**Datenübermittlung:
1.) „CDA-Dokument“
2.) „FHIR-Standard“**

Derzeit sind die mHealth/DiGA Anwendungen „Eco-Lösungen“ (Insellösungen wie z. B. „Herzmobil“) und müssen vornehmlich mit der Arztsoftware kompatibel sein bzw. bieten eigene Benutzeroberflächen für den GDA an.

**CDA Dokument:
Übergangsmittel
FHIR-Standard:
Übertragung von
Einzelwerten, noch nicht
in ELGA integriert**

Für eine ELGA-Kompatibilität von mHealth/DiGA gibt es grundsätzlich zwei Varianten der Datenübermittlung:

- **Variante 1:** Bis die neuen Schnittstellen, auf Basis des „Fast Healthcare Interoperability Resources“ (FHIR)-Standards [34], die die Übermittlung von Einzelwerten (z. B. Blutdruckwert, Gewicht) unterstützen, in ELGA implementiert sind, besteht die Möglichkeit von Anbietern von mHealth/DiGA ihre Einzelwerte gesammelt in Form eines Dokuments in ELGA einzubringen und damit mit allen zugriffsberechtigten ELGA-GDAs zu teilen. Konkret bedeutet das, dass aus den ermittelten Einzelwerten der mHealth/DiGA ein Episodenbericht in Form eines CDA-Dokuments entsteht. Der Episodenbericht kann entweder nach Zeit (täglich, monatlich, usw.) oder nach Ereignis (Eintreten eines Zwischenfalls oder Behandlungsziels) erstellt werden. Vorteil: einfache Integration in bestehende ELGA-kompatible Softwaresysteme (der CDA-Episodenbericht kann damit grundsätzlich von jedem ELGA-GDA gelesen werden). Nachteil: CDA-Dokumente werden meist nur als solche dargestellt, eine Weiterverarbeitung der Werte durch Praxis-/Kliniksoftware ist im Regelfall nicht vorgesehen. Auch

andere mHealth/DiGA Anwendungen können diese CDA-Dokumente meist nicht weiterverarbeiten.

- **Variante 2:** Übermittlung der Daten über den FHIR-Standard, der Schnittstellen auf Datenebene (Einzelwerte) bereitstellt. An diese Variante besteht eine hohe (internationale) Erwartungshaltung. Geplant ist zudem, dass mHealth/DiGA Anwendungen irgendeine Form von Zertifizierung/Qualitätssiegel erlangen sollen, bevor sie über FHIR-Schnittstellen Daten an ELGA senden können (insb. zur Sicherstellung der technischen und semantischen Kompatibilität, Einhalten von Prozessvorgaben und Vermeidung einer Datenflut). Erste Demos auf Basis der FHIR-Schnittstellen sind in 1-2 Jahren geplant. Vorteil: Drittsoftware (Praxis-/Kliniksoftware sowie andere mHealth/DiGA Anwendungen) kann diese Daten weiterverwenden. Nachteil: die Schnittstellen sind noch nicht vorhanden, Demos noch ausständig. Voraussetzung für Variante 2 ist, dass die mHealth/DiGA Anwendungen entsprechend validiert/bewertet wurden (z. B. diagnostische Genauigkeit), wenn mehrere Anwender die Daten weiterverwenden.

Exkurs: ELGA-Kompatibilität

Die Gesamtarchitektur von ELGA ist komplex, in Dokumenten auf der ELGA Website finden sich ausführliche Unterlagen dazu³. In der jetzigen Form ist ELGA ein geschlossenes System, welches für mHealth/DiGA noch nicht ausreichend ausgelegt ist. Im folgenden Kapitel soll eine Übersicht zum derzeitigen Stand von ELGA und den geplanten Entwicklungen und Erweiterungen gegeben werden. Außerdem sollen Aspekte und Voraussetzungen der ELGA-Kompatibilität für DiGA beschrieben werden.

**ELGA-Kompatibilität:
komplexe
Voraussetzungen**

Identifikation und Zugriffsberechtigungen in ELGA:

Derzeit können nur ELGA-Gesundheitsdiensteanbieter („ELGA-GDA“) ELGA für einen schreibenden Zugriff verwenden. Die Identifikation und Authentifizierung im niedergelassenen Bereich ist zudem an das e-card System gekoppelt (Patient*innen-Identifikation mittels e-card, Authentifizierung der ELGA-GDA mittels a-card).

- Apotheken sind erst im Zuge der Ausrollung der eMedikation mit einem e-card Zugang für ELGA ausgestattet worden.
- Eine Krankenanstalt fungiert beim Einloggen in ELGA als eine GDA-Einheit. Die Sicherstellung differenzierter interner Zugangsberechtigungen obliegt der Verantwortung der Krankenanstalt.
- Seit Einführung des eImpfpasses funktioniert als Identifikation (ausschließlich für den eImpfpass) auch der Login mittels Handy-Signatur bzw. zukünftig e-ID auf Basis der „Digitales Amt“ App des Bundesministeriums für Digitalisierung und Wirtschaftsstandort.
- Für die Ausweitung für weiterer, derzeit noch nicht an ELGA angebundener ELGA-GDA (insb. Wahlärzt*innen) wird gearbeitet.

³ Weiterführende Links: Dokument zur ELGA Gesamtarchitektur [35], Dokument zu den ELGA Schulungsunterlagen für GDA und SW-Hersteller [36], CDA-Leitfaden – Guide für Telemonitoring und Episodenberichte [37]

Der Ausbau des rollenbasierten Zugriffsmanagements ist für die zukünftige Ausgestaltung von ELGA wesentlich (Einsichtnahme vs. Abrufen von Informationen, wechselseitiger Zugriff). Derzeit funktioniert ELGA für Bürger*innen mono-direktional (nur Einsichtnahme/lesen ist möglich).

Technische Standards zur Interoperabilität:

**FHIR Standard:
ab 2023, für Einzelwerte**

ELGA ist gemäß den IHE-Konzepten („Integrating the Healthcare Enterprise“) aufgebaut. Diese sind spezialisiert für große Software-Anwendungen, jedoch nur bedingt geeignet für mHealth/DiGA.

Zukünftig soll daher auch der FHIR Standard („Fast Healthcare Interoperability Resources“) integriert werden um leichtgewichtige Schnittstellen für mHealth/DiGAs anbieten zu können. Mit dem FHIR-Standard ist es zudem möglich, einzelne, klar definierte Datenelemente, sogenannte „Ressourcen“ (wie z. B. Blutdruckwerte, Herzfrequenz, Gewicht) zu übertragen. Damit stellt dieser Standard auch die Interoperabilität zwischen mHealth/DiGA Anwendungen her (wurde beispielsweise ein Blutdruckwert von einer mHealth/DiGA Anwendung erhoben, können weitere mHealth/DiGA Anwendungen diesen Wert verwenden, ohne nochmals messen zu müssen). An der Integration von FHIR in der ELGA-Infrastruktur wird derzeit gearbeitet, eine erste Pilotierung wird voraussichtlich 2023 möglich sein.

Dokumente in ELGA:

**ELGA Mindestanforderung:
CDA-Dokumente**

ELGA arbeitet bislang noch nicht mit diesen Einzelwerten, sondern ist dokumentorientiert (CDA-Dokumente) aufgebaut. Aktuell verfügbar sind Entlassungsbriefe, Radiologiebefunde, Laborbefunde sowie verordnete und abgegebene Medikamente (e-Medikation). Der Ambulanzbefund und der sog. Episodenberichte werden mit der kommenden ELGA-Verordnung ermöglicht werden. Der detaillierte Aufbau der CDA-Dokumente wird mittels Implementierungsleitfäden normiert.

Weitere Voraussetzungen für die Kompatibilität mit ELGA:

**außerdem:
Identifizier (Patient*in und
Leistungserbringer),
Dokumentstruktur,
geschlossenes Netzwerk**

1. **Patient-ID:** Als digitaler Identifizier für Patient*innen wird das „Bereichsspezifische Personenkennzeichen Gesundheit“ (bPK-GH) verwendet, diese ist von der Stammzahl abgeleitet aber von anderen Identifiern (z. B. Finanzdaten) abgekoppelt. Die bPK-GH Identifizier liegen im sog. Zentralen Patient*innen-Index (Z-PI).
2. **Gesundheitsdienstleister-ID:** Als Identifizier für GDA werden hierarchische Object Identifier (OID) verwendet. GDA bekommen nach der Anmeldung (z. B. über das e-card System, das Krankenhausinformationssystem (KIS) oder, für den eImpfpass, über die Handy-Signatur) eine eindeutige OID zugeteilt, die im GDA-Index hinterlegt sind.
3. **CDA-Dokumentstruktur:** Diese ist in CDA-Leitfäden definiert. Darin sind Struktur und Codierung/Semantik des Dokuments festgelegt (z. B. SNOMED, LOINC, ASP, ...)
4. **Vertrauenswürdiger Identity-Provider:** Die Authentifizierung in ELGA basiert auf externen Identity-Providern, die eine Vertrauensstellung mit dem zentralen ELGA System besitzen (z. B. das e-card System im niedergelassenen Bereich). Für mobile DiGAs sind die Mechanismen des e-Governments (Handy-Signatur bzw. e-ID) dafür vorgesehen.

- Der Datenverkehr in ELGA wird aus Sicherheitsgründen derzeit nicht über das Internet, sondern über eigene geschlossene Gesundheitsnetzwerke geführt. Mit der Öffnung von ELGA in Richtung mHealth/DiGA ist auch die Erreichbarkeit für mobile Applikationen sicherzustellen.

Eine mHealth/DiGA muss also das Konzept der bPK-GH (für die Identifikation von Patient*inne) und der OID (für die Identifikation von ELGA-GDA) sowie die Authentifizierung mittels vertrauenswürdigen Identity-Provider unterstützen, um mit ELGA kompatibel zu sein. Bereits jetzt besteht grundsätzlich die Möglichkeit für mHealth/DiGA Anwendungen, sich am Z-PI anzubinden, um die bPK-GH als Identifikator zu beziehen.

Eine Veranschaulichung über die technische Kompatibilität von DiGAs mit ELGA, basierend auf der Expert*innenkonsultation mit der ELGA GmbH, ist in der nachfolgenden Abbildung 3-2 dargestellt.

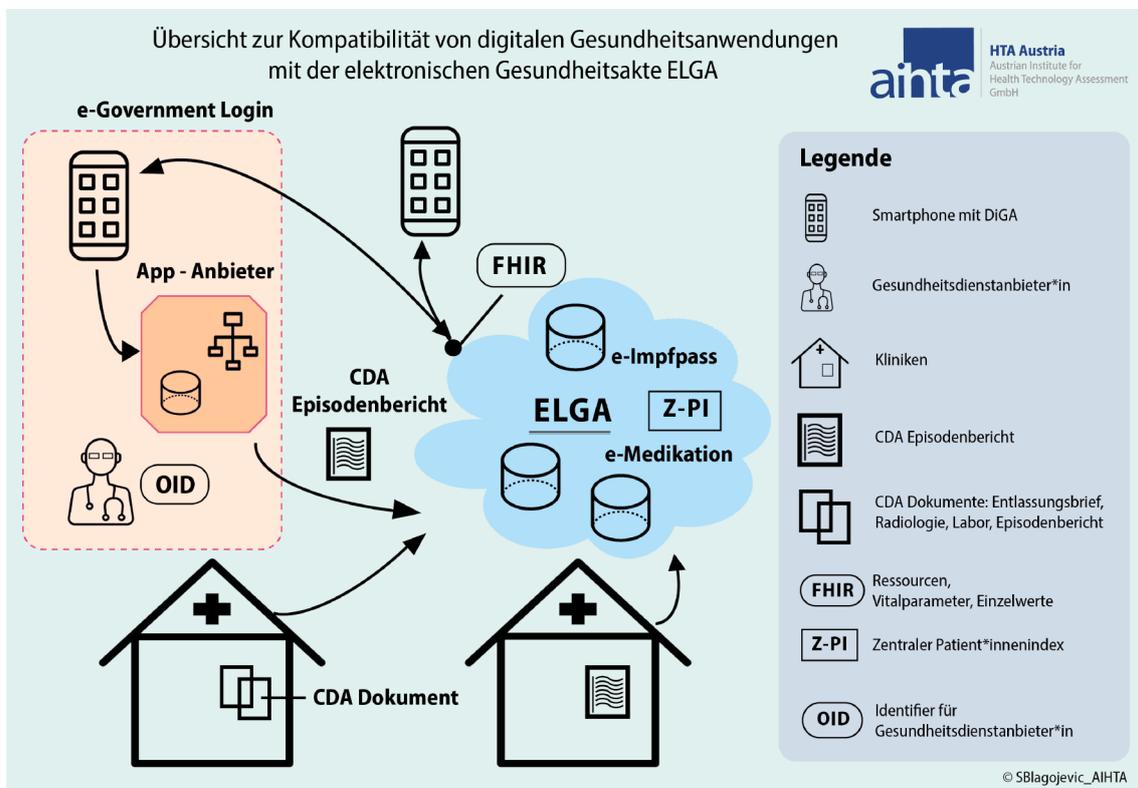


Abbildung 3-2: Übersicht der Kompatibilität von digitalen Gesundheitsanwendungen mit der elektronischen Gesundheitsakte ELGA, basierend auf der Expert*innenkonsultation mit der ELGA GmbH

3.2.2 Überblick über technologische Aspekte der DiGAs mit ASVG-Relevanz

Für das Verfahren der strukturierten und evidenzbasierten Bewertung von DiGAs spielen die technologie-spezifischen Aspekte eine wesentliche Rolle. Um eine Aussage über das Vorhandensein verschiedener technologie-spezifischer Charakteristika zu treffen, folgt eine exemplarische Analyse für die im vorigen Kapitel nach den Priorisierungskriterien ausgewählten DiGAs.

**Beschreibung
priorisierter DiGAs
in 4 Gruppen nach
Indikation**

Im Folgenden wird zunächst eine detailliertere Beschreibung der DiGAs mit ASVG-Relevanz und deren Anwender*innen, Zielgruppe, Funktion sowie Indikationen gegeben. Für eine bessere Übersicht werden sie in folgende Indikationsgebiete gruppiert beschrieben:

- Herz-Kreislaufkrankungen
- Endokrine, Ernährungs- und Stoffwechselerkrankungen
- Psychische und Verhaltensstörungen
- Andere

Nach der Beschreibung der DiGAs mit ASVG-Relevanz folgt in Kapitel 3.2.3 zur beispielhaften Veranschaulichung eine Übersicht über die technologie-spezifischen Aspekte: Risikoklasse, Datenschutz, Algorithmen, verwendete technologische Standards für Interoperabilität sowie ELGA-Kompatibilität (= Verwendung des „HL7 CDA-Standard“), vgl. Tabelle 3-7.

Herz-Kreislaufkrankungen

**7 DiGAs für Herz-
Kreislaufkrankungen**

Es wurden sieben DiGAs für Herz-Kreislaufkrankungen gefunden. Die Behandlung unterstützt vor allem die Diagnosen Herzinsuffizienz, Herzrhythmusstörungen/Vorhofflimmern, chronisch obstruktive Lungenerkrankung und zerebrovaskuläre Erkrankungen (vgl. Tabelle 3-3).

CardiacSense (Arseus Hospital NV) [38]

CardiacSense

Die CardiacSense-Anwendung zeigt den Nutzer*innen ein Echtzeit-Daten-Dashboard, erstellt Ereignisberichte und teilt diese gemeinsam mit den Vitalparametern mit den behandelnden Ärzt*innen. Die Messungen werden von den Sensoren der CardiacSense Watch durchgeführt, die in Kontakt mit der Haut des Handgelenks der Patient*innen ist.

CardioMemory [39]

CardioMemory

CardioMemory ist ein mobiles Gesundheitsdaten-Tagebuch, mit welchem relevante Gesundheitsdaten (Blutdruck, Herzfrequenz, Schritte pro Tag, Körpergewicht, Wohlbefinden, Medikation und allgemeine Kommentare) mit der Applikation erfasst und als tabellarisches Tagebuch dargestellt werden.

FibriCheck (Qompium) [40]

FibriCheck

Die Anwendung FibriCheck ist ausschließlich für die Aufzeichnung, Anzeige, Speicherung und Übertragung von Photoplethysmogramm-Daten (Aufzeichnung der Pulswellenform zur Überwachung des Herzrhythmus) bestimmt. Die Aufzeichnungen erfolgen durch Kontakt zwischen Haut der Fingerkuppe und Kamera eines mobilen Geräts oder Wearable-Geräts. Auf diese Weise können volumetrische Veränderungen erfasst werden.

Heartfish [41]

Die App bietet Funktionen zur Dokumentation von Trainings, personalisierte Trainingsempfehlungen, Bluetooth-gestützte Steuerung des Trainings, Vitalparameterdokumentation, Fortschrittvizualisierung, integriertes Coachingsystem und PDF-Berichte.

Heartfish

HerzMobil [39]

HerzMobil ist ein mobiles Gesundheitsdaten-Tagebuch, entwickelt vom Austrian Institute for Technology (AIT). Es kann in das webbasierte Therapie Management System „HerzMobil integriert werden. So können relevante Gesundheitsdaten (Blutdruck, Herzfrequenz, Körpergewicht, Wohlbefinden, Medikation) mit der Applikation erfasst und als tabellarisches Tagebuch dargestellt werden. Messdaten können über NFC fähige Messgeräte automatisch ausgelesen werden (unterstützte Standards: NDEF, FeliCa). Eine Freischaltung und Benutzung der Online-Version ist auf Partner der Tirol Kliniken beschränkt. Diese umfasst zusätzlich zur oben beschriebenen Funktionalität, die automatische Synchronisation mit dem individuellen Online-Tagebuch und den Empfang von Feedback-Information durch den betreuenden Arzt.

HerzMobil

Rehappy (Rehappy GmbH) [42]

Rehappy unterstützt die Nachsorge von Schlaganfallpatient*innen mittels individuell zusammengestellten Motivations- und Wissensversorgungen, einem Aktivitätstracker und einem Web-Portal. Die Begleitung basiert auf edukativer Informationsvermittlung und positiver Verstärkung zur Wahrnehmung der Eigenverantwortung und Steigerung der Therapieadhärenz sowie intrinsischen Motivation.

Rehappy

MobECG (Savvy) [43]

Die Applikation MobECG bietet eine EKG-Überwachung von Vorhofflimmern zur Vorbeugung von kryptogenen Schlaganfällen oder transienten ischämischen Attacken.

MobECG

Tabelle 3-3: Kategorisierung von DiGAs: Herz-Kreislauferkrankungen

Name der App (Hersteller)	Plattformen: Verfügbarkeit in Google Playstore/Apple App Store (J/N)	Verzeichnis	Anwender*innen	Zielgruppe	Funktion der DiGA	Patient*innengruppe
CardiacSense (Arseus Hospital NV) [38]	Webanwendung	BE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring	<ul style="list-style-type: none"> ■ Herz- und Lungenerkrankungen <ul style="list-style-type: none"> ■ Vorhofflimmern ■ chronisch obstruktive Lungenkrankheit <ul style="list-style-type: none"> ■ Herzinsuffizienz ■ Schlafapnoe ■ Schlaganfallrisiko
CardioMemory [39]	App (J/N)	AT – AIT; AT – ifGP	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Herz-Kreislauferkrankungen
FibriCheck (Qompium) [40]	Web and App	BE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring	<ul style="list-style-type: none"> ■ Herzerkrankungen ■ Vorhofflimmern
Heartfish [41]	App (J/J)	AT – ScaleUp4Europe	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Herz-Kreislauferkrankungen
HerzMobil [39]	App (J/N)	AT – AIT; AT – ifGP	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Herz-Kreislauferkrankungen
Rehappy (Rehappy GmbH) [42]	App (J/J)	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Zerebrale transitorische Ischämie und verwandte Syndrome <ul style="list-style-type: none"> ■ Subarachnoidalblutung ■ Intrazerebrale Blutung ■ Sonstige nichttraumatische intrakranielle Blutung <ul style="list-style-type: none"> ■ Hirninfarkt ■ Schlaganfall, nicht als Blutung oder Infarkt bezeichnet <ul style="list-style-type: none"> ■ Sonstige zerebrovaskuläre Krankheiten ■ Folgen einer zerebrovaskulären Krankheit
MobECG (Savvy) [43]	App (J/J)	AT – ScaleUp4Europe	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring	<ul style="list-style-type: none"> ■ Vorhofflimmern

Legende Verzeichnisse: AT: Österreich (AIT: Austrian Institute of Technology; IfGP: Institut für Gesundheitsförderung und Prävention); BE: Belgien (mHealth Belgium Platform); DE: Deutschland (DiGA-Verzeichnis); DiGA – Digitale Gesundheitsanwendung; UK: United Kingdom (National Health Service App Library);

Endokrine, Ernährungs- und Stoffwechselerkrankungen

Es wurden sieben DiGAs für Endokrine, Ernährungs- und Stoffwechselerkrankungen gefunden. Die Behandlung unterstützt vor allem die Diagnosen Diabetes und Adipositas (vgl. Tabelle 3-4).

7 DiGAs für Endokrine, Ernährungs- und Stoffwechselerkrankungen

DiabCare [39]

DiabCare ist ein mobiles Diabetes Gesundheitsdaten-Tagebuch (entwickelt vom AIT), das im Rahmen eines Forschungsprojektes mit dem webbasierten Therapie Management System „DiabCare“ verwendet werden kann. Relevante Gesundheitsdaten (Blutzucker, Blutdruck, Herzfrequenz, Schritte/Tag, Körpergewicht, Wohlbefinden, Ereignisse und Freitextnotizen) können mit der Applikation erfasst und als tabellarisches Tagebuch dargestellt werden.

DiabCare

DiabMemory 2 [39]

DiabMemory 2 ist ein mobiles Diabetes Gesundheitsdaten-Tagebuch, welches vom AIT – im Rahmen eines Forschungsprojektes der Versicherungsanstalt öffentlich Bediensteter, Eisenbahnen und Bergbau (BVAEB) – entwickelt wurde. DiabMemory 2 kann in das webbasierte Therapie Management System „DiabMemory“ integriert werden.

DiabMemory 2

Freestyle LibreLink – BE (Abbott SA/NV) [44]

Freestyle LibreLink kann den aktuellen Blutzuckermesswert und den Blutzuckerverlauf anzeigen. Außerdem können Notizen für Ernährung, Insulinverwendung und Bewegung erstellt werden. Mit LibreView und LibreLinkUp ist eine Schnittstelle zu Gesundheitspersonal möglich, um Glukoseberichte anzuzeigen.

Freestyle LibreLink

Guardian Connect App (Medtronic Belgium) [45]

Die Guardian™ Connect-Anwendung ist eine Komponente des Guardian Connect Systems zur kontinuierlichen Glukoseüberwachung um die Diabetes-Behandlung zu unterstützen. Die Funktionen dieser DiGA sind die kontinuierliche Aufzeichnung der Glukosewerte, Anzeige der Glukosewerte, Benachrichtigungen bei Glukose-Entgleisungen, Anzeige der Auswirkungen von Ernährung, Bewegung und Medikamenten auf die Glukosewerte.

Guardian Connect App

mySugr [46]

mySugr ist eine Diabetes-App mit digitalem Logbuch. Damit kann der Blutzucker verfolgt, Ihre Kohlenhydratzufuhr aufgezeichnet und der HbA1c-Wert geschätzt werden.

mySugr

Vitadio [47]

Vitadio Health ist ein digitales Therapieprogramm, das Patient*innen bei der Prävention und Behandlung von Typ-2-Diabetes durch verhaltenstherapeutische Unterstützung.

Vitadio

zanadio (aidhere GmbH) [48]

zanadio ist eine Anwendung zur Gewichtsreduktion für Patient*innen mit Adipositas. Durch eine Veränderung der Verhaltensweisen (Ernährung und Bewegung) sollen die Patient*innen unterstützt werden. Die DiGA basiert dabei auf dem wissenschaftlichen Konzept der multimodalen, konservativen Adipositas-Therapie, welche die verschiedenen relevanten Bereiche adressiert und hierdurch eine langfristige, dauerhafte Gewichtsreduktion herbeiführt.

zanadio

Tabelle 3-4: Kategorisierung von DiGAs: Endokrine, Ernährungs- und Stoffwechselerkrankungen

Name der App (Hersteller)	Plattformen: Verfügbarkeit in Google Playstore/Apple App Store (J/N)	Verzeichnis	Anwender*innen	Zielgruppe	Funktion der DiGA	Patient*innengruppe
DiabCare [39]	App (J/N)	AT – AIT	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	■ Diabetes Mellitus
DiabMemory 2 [39]	App (J/J)	AT – AIT	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	■ Diabetes Mellitus
Freestyle LibreLink – BE (Abbott SA/NV) [44]	App (J/J)	BE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring	■ Diabetes Mellitus
Guardian Connect App (Medtronic Belgium) [45]	App (J/J)	BE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring	■ Diabetes Mellitus
mySugr Tagebuch-App [46]	App (J/J)	AT – ifGP; BE; E	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	■ Diabetes Mellitus
Vitadio [47]	App (J/J)	AT – ScaleUp4Europe	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	■ Diabetes Mellitus
zanadio (aidhere GmbH) [48]	App (J/J)	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	■ Adipositas

Legende Verzeichnisse: AT: Österreich (AIT: Austrian Institute of Technology; IfGP: Institut für Gesundheitsförderung und Prävention); BE: Belgien (mHealth Belgium Platform); DE: Deutschland (DiGA-Verzeichnis); UK: United Kingdom (National Health Service App Library); Abkürzungen: DiGA – Digitale Gesundheitsanwendung

Psychische Erkrankungen

Es wurden neun DiGAs für psychische Erkrankungen gefunden. Die Behandlung unterstützt vor allem die Diagnosen Depression, Angststörungen und Schlafstörungen (vgl. Tabelle 3-5).

9 DiGAs für psychische Erkrankungen

Catch It [49]

Catch It ist eine DiGA, um einen besseren Umgang mit Gefühlen der Angst und Depression zu erlernen. Mit der App soll vermittelt werden, Probleme anders zu betrachten, negative Gedanken in positive zu verwandeln und das psychische Wohlbefinden zu verbessern.

Catch It

deprexis (GAIA AG) [50]

deprexis ist ein Selbsthilfeprogramm zur Therapieunterstützung von Patient*innen mit Depressionen und depressiven Verstimmungen, die mindestens 18 Jahre alt sind. deprexis basiert auf etablierten psychotherapeutischen Ansätzen und Verfahren (insbesondere der kognitiven Verhaltenstherapie).

deprexis

Invirto (Sympatient GmbH) [51]

Invirto kann als Zusatz zur Verhaltenstherapie verwendet werden: Im Kern steht dabei die Simulation der (Angst auslösenden) Situationen mittels Virtual-Reality-Brille. Anschließend können Übungen auf einer App durchgeführt werden.

Invirto

Kalmeda (mynoise GmbH) [52]

Kalmeda bietet Patient*innen mit chronischer Tinnitus-Belastung eine leitlinienbasierte, verhaltenstherapeutische Therapie. Zusätzlich gibt es Entspannungsanleitungen, beruhigende Natur- und Hintergrundgeräusche sowie Informationen zur Diagnose zur Patient*innenedukation.

Kalmeda

Mindable (Mindable Health GmbH) [53]

Mindable unterstützt Nutzer*innen mit Angsterkrankungen. Das Behandlungsprogramm basiert auf S3-leitlinienkonformen Methoden der Kognitiven Verhaltenstherapie. Neben Psychoedukation liegt der Schwerpunkt des Behandlungsprogramms auf der Gewöhnung an Körpersymptome sowie in-vivo Konfrontation. Während der Konfrontationsübungen werden Angstverläufe live aufgezeichnet. Darüber hinaus können Nutzer*innen angstbezogene Symptome, Ereignisse und Verhalten erfassen und bei Bedarf mit den behandelnden Arzt*innen oder Therapeut*innen teilen.

Mindable

Selfapy Online-Kurs bei Depressionen (Selfapy GmbH) [54]

Selfapy ist eine DiGA für Patient*innen mit Depression. Mit Selfapy wird ein individueller Online-Kurs angeboten, der auf evidenzbasierten Theorien und Techniken der Kognitiven Verhaltenstherapie basiert. Dabei wird der Umgang mit negativen Gedanken, der Erstellung einer positiven Tagesstruktur, Entspannungstechniken, Schlafproblemen, sowie Strategien zur Rückfallprävention erarbeitet. Die Patient*innen werden während des Kurses durchgehend von Psycholog*innen begleitet, bei akutem Bedarf können die Psycholog*innen über eine Nachrichtenfunktion kontaktiert werden.

Selfapy

somnio (mementor DE GmbH) [55]

somnio somnio ist eine digitale Anwendung zur Behandlung von Ein- und Durchschlafstörungen (Insomnie). Dabei wird die evidenzbasierte und Leitlinienkonforme kognitive Verhaltenstherapie für Insomnie verwendet. Nutzer*innen lernen die Schlafzeiten zu optimieren, einem individuell abgestimmten Schlaf-Wach-Rhythmus zu folgen, mit schlafhindernden Gedanken umzugehen oder sich mittels Entspannungstechniken in einen schlafförderlichen Zustand zu bringen.

velibra (GAIA AG) [56]

velibra velibra ist eine DiGA für Patient*innen mit Angststörungen. Die DiGA vermittelt etablierte Methoden und Übungen der Kognitiven Verhaltenstherapie. Das Programm ist als Ergänzung zu einer sonst üblichen Behandlung für Patient*innen vorgesehen.

vorvida (GAIA AG) [57]

vorvida vorvida ist eine DiGA für Patient*innen mit einer psychischen- und Verhaltensstörung durch Alkohol. Die Patient*innen sollen beim Management des gesundheitsschädlichen Trinkverhaltens unterstützt werden um die Trinkmenge zu reduzieren. Die DiGA basiert auf psychotherapeutischen Ansätzen und Verfahren der kognitiven Verhaltenstherapie. Das Programm soll in Ergänzung zu einer sonst üblichen Behandlung eingesetzt werden.

Tabelle 3-5: Kategorisierung von DiGAs: psychische Erkrankungen

Name der App (Hersteller)	Plattformen: Verfügbarkeit in Google Playstore/Apple App Store (J/N)	Verzeichnis	Anwender*innen	Zielgruppe	Funktion der DiGA	Patient*innengruppe
Catch It [49]	App (J/J)	UK	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Angstzustände ■ Depression
Deprexis (GAIA AG) [50]	Webanwendung	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Depression ■ depressive Episode
Invirto (Sympatient GmbH) [51]	App (J/J)	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	8 Direkte Intervention	<ul style="list-style-type: none"> ■ Agrarphobie ■ Panikstörungen ■ soziale Phobien
Kalmeda (mynoise GmbH) [52]	App (J/J)	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	8 Direkte Intervention	<ul style="list-style-type: none"> ■ Tinnitus
Mindable: Panikstörungen und Agoraphobie (Mindable Health GmbH) [53]	App (J/J)	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	8 Direkte Intervention	<ul style="list-style-type: none"> ■ Agrarphobie ■ Panikstörung
Selfapys Online-Kurs bei Depressionen (Selfapy GmbH) [54]	Webanwendung	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Depression ■ depressive Episode
somnio (mementor DE GmbH) [55]	Web and App	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Nichtorganische Insomnie
velibra (GAIA AG) [56]	Webanwendung	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Agrarphobie ■ soziale Phobie ■ Panikstörungen ■ generalisierte Angststörung
vorvida (GAIA AG) [57]	Webanwendung	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Alkohol: Schädlicher Gebrauch ■ Alkohol: Abhängigkeitssyndrom

Legende Verzeichnisse: AT: Österreich (AIT: Austrian Institute of Technology; IfGP: Institut für Gesundheitsförderung und Prävention); BE: Belgien (mHealth Belgium Platform); DE: Deutschland (DiGA-Verzeichnis); UK: United Kingdom (National Health Service App Library); Abkürzungen: DiGA – Digitale Gesundheitsanwendung

Digitale Gesundheitsanwendungen für andere Erkrankungen

15 DiGAs für andere Indikationen	<p>Es wurden 15 DiGAs für sonstige Indikationen gefunden. Diese Indikationen umfassten bösartige Neubildungen, neurologische Erkrankungen, orthopädische Erkrankungen, hämatologischen Erkrankungen sowie unspezifische chronische Erkrankungen oder Patient*innen mit Dauermedikation (vgl. Tabelle 3-6).</p>
	<p><i>AirviewTM (Resmed) [58]</i></p>
AirviewTM	<p>AirView ist eine cloudbasierte Technologie, die für Patient*innenscreening, Diagnose und Therapiemanagement von Schlaf- und Beatmungspatient*innen verwendet werden kann. Es soll dabei ein zeitnaher Fernzugriff von Gesundheitspersonal auf Therapie- und Geräteinformationen ermöglicht werden.</p>
	<p><i>CANKADO PRO-React Onco (CANKADO Service GmbH) [59]</i></p>
CANKADO PRO-React Onco	<p>CANKADO PRO-React Onco ist eine DiGA für Patient*innen mit bösartigen Neubildungen der Brustdrüse. In Abhängigkeit zur Erkrankung und der laufenden Therapie werden automatisiert Verhaltenshinweise gegeben, wie dringend die jeweilige Beschwerde mit den behandelnden Ärzt*innen besprochen werden sollte.</p>
	<p><i>Comarch Home Health (Comarch) [60]</i></p>
Comarch Home Health	<p>Comarch Home Health ist eine DiGA für Patient*innen mit chronischen Erkrankungen. Die DiGA beinhaltet Funktionen zur Aufzeichnung der Messungen und Fragebögen, um sie beim Besuch bei Ärzt*innen vorlegen zu können. Die persönlichen Daten werden in einer Telemedizinplattform gespeichert. Wenn Schwellenwerte überschritten werden, sendet die DiGA Informationen an die Plattform. Eine Warnung wird den Ärzt*innen angezeigt und ermöglicht eine Reaktion durch Videokonsultation.</p>
	<p><i>Comunicare (Comunicare Solutions) [61]</i></p>
Comunicare	<p>Comunicare ist eine DiGA für Patient*innen mit chronischen Erkrankungen. Die DiGA wird zur Stärkung der Eigenverantwortung der Patient*innen, der Lesekompetenz, der Therapietreue und der Selbstwirksamkeit eingesetzt. Die Patient*innen zur Verfügung gestellten Informationen werden von medizinischen Expert*innen erstellt und durch Algorithmen werden personalisierte Informationen über Pathologie und Therapie angezeigt. Patient*innen können physiologische Parameter, Symptome und Fragen zur Lebensqualität dokumentieren. Das Gesundheitspersonal kann die Patient*innendaten mittels Dashboard verwalten und Telekonsultationen durchführen.</p>
	<p><i>Elevida (GAIA AG) [62]</i></p>
Elevida	<p>Elevida ist eine DiGA für Patient*innen mit multipler Sklerose und gleichzeitiger Fatigue. Dabei werden evidenzbasierte psychotherapeutische Verfahren eingesetzt, um die Patient*innen beim Management der Fatigue zu unterstützen.</p>
	<p><i>Florio (Swedish Orphan Biovitrum (Belgium) BV) [63]</i></p>
Florio	<p>Florio ITP ist eine DiGA mit digitalen Tagebuchfunktionen. Mit dieser DiGA können Thrombozytenwerte und Symptome dokumentiert werden. Außerdem können Erinnerungen zur Medikamenteneinnahme gesetzt werden.</p>

HumanITcare [64]

HumanITCare ist eine telemedizinische Anwendung, die es Ärzt*innen ermöglicht, den Überblick über ihre Patient*innen zu behalten.

HumanITcare*icompanion (icometrix) [65]*

icompanion ist eine DiGA für Patient*innen mit multipler Sklerose. Die DiGA beinhaltet Funktionen zur Dokumentation der täglichen Stimmung, Behandlungen, Symptome, Therapien und Tagebuchnotizen durch Patient*innen. Außerdem können Tests zur Bewertung des Grades der Behinderung, der kognitiven Fähigkeiten und der Müdigkeit durchgeführt werden. Patient*innen können zusätzlich MRT-Scans hochladen und ansehen.

icompanion*Mika (Fosanis GmbH) [66]*

Mika ist eine DiGA für Patient*innen mit Krebserkrankungen. Funktionen: Dokumentation von Belastungen, Symptomen und Nebenwirkungen im Verlauf. Außerdem Patientinnen-Edukation in den Bereichen Gesundheitskompetenz, Stressmanagement, Bewegung und Ernährung.

Mika*Mindmate [67]*

MindMate ist eine DiGA für Patient*innen mit Demenz. Ratschläge zur richtigen Ernährung und körperlichen Bewegung sollen die Gehirnaktivität stimulieren und helfen, mit dem sozialen Umfeld in Verbindung zu bleiben. Die DiGA beinhaltet zusätzlich Engagement-Tools wie eine Jukebox. Diese Tools sollen Patient*innen mit Demenz besser einbinden, und Pflegepersonal eine personenzentrierte Pflege ermöglichen.

Mindmate*M-sense (Newsenselab GmbH) [68]*

M-sense bietet ein digitales Behandlungsprogramm für Patient*innen mit Migräne. Die Anwendung beinhaltet ein digitales Kopfschmerztagebuch sowie leitlinienkonforme Verfahren zur Migräneprophylaxe und Akutbehandlung. Die Funktionen der DiGA beinhalten Wissensvermittlung, animierte physiotherapeutische Übungen, Anleitungen für Ausdauersport, Entspannungs- und Imaginationenübungen. Lebensstilfaktoren können im Tagebuch dokumentiert werden.

M-sense*Noona (Varian Medical Systems Belgium) [69]*

Noona ist eine DiGA für Patient*innen mit Krebserkrankungen sowie Pflegende. Die Funktionen der DiGA beinhalten Kommunikationswerkzeuge mit cloudbasierten „patient reported outcomes“. Symptome sollen in Echtzeit erfasst und strukturiert dem Pflorgeteam präsentiert werden. Dadurch können Patient*innen auf Basis der Symptomstärke priorisiert werden.

Noona*Plateforme Maela (Medtronic Belgium) [70]*

Plateforme Maela ist eine DiGA für die Einbindung, Aufklärung und Überwachung von Patient*innen mit Indikationen für eine Operation. Über mehrere Kanäle (Web, Mobil, SMS) können Symptome registriert, Vitalparameter eingegeben und relevantes Informationsmaterial eingesehen werden. Die DiGA bietet einen Kommunikationskanal mit dem Pflorgeteam.

Plateforme Maela

Vaica [71]

Vaica Vaica ist eine DiGA für Patient*innen mit Dauermedikation. Die DiGA dient dem Patient*innenmonitoring und unterstützt die Adhärenz. Vaica kombiniert automatische Medikamentenspender („smart medication dispensers“) mit cloudbasierter Software.

Vivira (Vivira Health Lab GmbH) [72]

Vivira Vivira ist eine DiGA zur Behandlung von orthopädischen Schmerzen. Die Funktionen der DiGA sind täglich vier bewegungstherapeutische Übungen, die nach Rückmeldungen der Patient*innen fortlaufend angepasst werden. Die Trainings-Elemente basieren auf Leitlinienempfehlungen. Weitere Funktionen sind Abfragen zur Gesundheit, Visualisierung des Fortschritts, Bewegungstests und edukative Inhalte.

Tabelle 3-6: Kategorisierung von DiGAs: andere Indikationen

Name der App (Hersteller)	Plattformen: Verfügbarkeit in Google Playstore/Apple App Store (J/N)	Verzeichnis	Anwender*innen	Zielgruppe	Funktion der DiGA	Patient*innengruppe
Comarch Home Health (Comarch) [60]	App (J/J)	BE	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring	<ul style="list-style-type: none"> ■ chronischen Krankheiten ■ ältere Menschen ■ nach einem Krankenhausaufenthalt ■ während des Diagnoseprozesses
Comunicare (Comunicare Solutions) [61]	App (J/J)	BE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring	<ul style="list-style-type: none"> ■ chronische Krankheiten
HumanITcare [64]	App (J/J)	AT – ScaleUp4-Europe	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring	<ul style="list-style-type: none"> ■ Chronische oder infektiösen Erkrankungen
Plateforme Maela (Medtronic Belgium) [70]	Web and App	BE	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring	<ul style="list-style-type: none"> ■ Während Therapien (z. B. kolorektale, bariatrische, Wirbelsäulen-, Thoraxchirurgie)
Vaica [71]	Web and App	AT – HealthHub-Vienna	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Patient*innen unter medikamentöser Therapie
CANKADO PRO-React Onco (CANKADO Service GmbH) [59]	Web and App (J/J)	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Bösartige Neubildung der Brustdrüse
Mika (Fosanis GmbH) [66]	App (J/J)	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Bösartige Neubildung der Cervix uteri ■ Bösartige Neubildung des Uterus ■ Bösartige Neubildung des Ovars
Noona (Varian Medical Systems Belgium) [69]	App (J/J)	BE	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	5 Einfaches Monitoring	<ul style="list-style-type: none"> ■ Krebspatient*innen und Pflegende
AirviewTM (Resmed) [58]	App (J/J)	BE	Kliniker/Gesundheitspersonal und Gesundheitsinstitutionen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ schlafbezogener Atmungsstörung ■ schlafbezogener Atmungsausfall
Florio (Swedish Orphan Biovitrum (Belgium) BV) [63]	Web and App	BE	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	5 Einfaches Monitoring	<ul style="list-style-type: none"> ■ Hämophilie A oder B
Vivira (Vivira Health Lab GmbH) [72]	App (J/J)	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Koxarthrose ■ Gonarthrose Sonstige Gelenkrankheiten: <ul style="list-style-type: none"> ■ Osteochondrose der Wirbelsäule ■ Sonstige Krankheiten der Wirbelsäule und des Rückens <ul style="list-style-type: none"> ■ Rückenschmerzen ■ Biomechanische Funktionsstörungen
Elevida (GAIA AG) [62]	Webanwendung	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Multiple Sklerose
icompanion (icomatrix) [65]	Web and App	BE	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	5 Einfaches Monitoring	<ul style="list-style-type: none"> ■ Multiple Sklerose
M-sense Migräne (Newsenselab GmbH) [68]	App (J/J)	DE	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Migräne
Mindmate [67]	Web and App	AT – HealthHub-Vienna	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention	<ul style="list-style-type: none"> ■ Demenz

Legende Verzeichnisse: AT: Österreich (AIT: Austrian Institute of Technology; IfGP: Institut für Gesundheitsförderung und Prävention); BE: Belgien (mHealth Belgium Platform); DE: Deutschland (DiGA-Verzeichnis); UK: United Kingdom (National Health Service App Library); Abkürzungen: DiGA – Digitale Gesundheitsanwendung

3.2.3 Technologie-spezifische Aspekte: Risikoklassen, Datenschutz, Algorithmen, technische Standards für Interoperabilität, ELGA-Kompatibilität

**technische Charakteristika:
Prüfung vor
Nutzen-Evaluation**

Für die nach den in Kapitel 3.1 beschriebenen Priorisierungskriterien identifizierten DiGAs gibt es viele technische Charakteristika. Diese können entscheidend sein, ob eine DiGA überhaupt verwendet werden kann. Eine Auflistung der in diesem Abschnitt erhobenen technologie-spezifischen Charakteristika findet sich im Methodenkapitel 2.2.

**29 DiGAs:
CE-Kennzeichnung;
10 DiGAs:
Herstellerangabe zu
Risikoklasse;
35 DiGAs:
Datenschutz-Angaben**

Für die 38 priorisierten DiGAs zeigte sich, dass für 29 DiGAs eine CE-Kennzeichnung vorliegt. Bezüglich Risikoklasse nach Medizinproduktverordnung 2017/745 wurden drei DiGAs in Risikoklasse 1, sechs DiGAs in Risikoklasse 2a und eine DiGA in Risikoklasse 2b eingeteilt. Für 28 DiGAs wurde keine Angabe zur Risikoklasse laut Hersteller gemacht. Zum Datenschutz liegen für 35 DiGAs eine Datenschutzerklärung vor. Für zwei DiGAs gibt es keine Angaben zum Datenschutz, für eine DiGA wird ein Re-Assessment des Datenschutzes durchgeführt. Angaben zu verwendeten Algorithmen wurden für keine der priorisierten DiGAs identifiziert. Alle 38 priorisierten DiGAs sind nach Anwendung des NICE Evidence Standards Frameworks [11] in die höchste Evidenzstufe „C“ einzuteilen.

**14 DiGAs:
Angaben zu
technischen Standards**

Für 24 der priorisierten DiGAs wurden keine Angaben zu verwendeten technischen Standards für Interoperabilität gemacht. Für 14 der DiGAs wurden technischen Standards für Interoperabilität genannt (Details zu den jeweiligen Standards vgl. Tabelle 3-7). Insgesamt erfüllen sieben der DiGAs zum jetzigen Zeitpunkt die Voraussetzungen für eine ELGA-Kompatibilität (Verwendung des HL7 CDA Standards). Ob eine technische Kompatibilität mit ELGA ein zusätzliches Filterkriterium darstellen könnte, wird im Kapitel 3.5 diskutiert.

**7 DiGAs:
ELGA-Kompatibilität
(HL7 CDA Standard)**

In der nachfolgenden Tabelle 3-7 ist eine Übersicht über die technologiespezifischen Anforderungen der DiGAs, vorhandene Schnittstellen für Interoperabilität und ELGA-Kompatibilität, sowie deren Einteilung in Risikoklassen und Evidenzstufen gegeben. Dabei wird deutlich, dass eine umfassende Evaluation aufgrund fehlender Angaben (zur CE-Kennzeichnung, zur Risikoklasse, zum Datenschutz, zu den verwendeten Algorithmen, sowie zu technischen Standards für Interoperabilität) nicht immer möglich ist.

Tabelle 3-7: Technologiespezifische Anforderungen und Einteilung der DiGAs in Risikoklassen sowie Evidenzstufen

Name der App (Hersteller)	CE Kennzeichnung*	Risikoklasse laut Hersteller*	Risikoklasse nach 2017/745 MDR*	Datenschutz**	Algorithmen***	Evidenzstufe (NICE) [11]	Interoperabilität	ELGA-kompatible Technologie
AirviewTM (Resmed) [58]	✓	2A	2A	✓	k.A.	C	KMEHR, HL7 CDA, HL7 FHIR, SNOMED-CT	✓
CANKADO PRO-React Onco (CANKADO Service GmbH) [59]	✓	1	k.A.	✓	k.A.	C	GAMP-5, ICH GCP E6(R2) HL7 FHIR	Nein
CardiacSense (Arseus Hospital NV) [38]	✓	2A	2A	✓	k.A.	C	k.A.	k.A.
CardioMemory [39]	k.A.	k.A.	k.A.	✓	k.A.	C	k.A.	k.A.
Catch It [49]	k.A.	k.A.	k.A.	Derzeit Re-Assessment	k.A.	C	k.A.	k.A.
Comarch Home Health (Comarch) [60]	✓	2A	2A	k.A.	k.A.	C	k.A.	k.A.
Comunicare (Comunicare Solutions) [61]	✓	1	1	✓	k.A.	C	KMEHR, HL7 CDA, HL7 FHIR, SNOMED-CT,	✓
Deprexis (GAIA AG) [50]	✓	1	k.A.	✓	k.A.	C	HL7 FHIR	Nein
DiabCare [39]	k.A.	k.A.	k.A.	✓	k.A.	C	k.A.	k.A.
DiabMemory 2 [39]	k.A.	k.A.	k.A.	✓	k.A.	C	k.A.	k.A.
Elevida (GAIA AG) [62]	✓	1	k.A.	✓	k.A.	C	HL7 FHIR	Nein
FibriCheck (Qompium) [40]	✓	2A	2A	✓	k.A.	C	k.A.	k.A.
Florio (Swedish Orphan Biovitrum (Belgium) BV) [63]	✓	1	1	✓	k.A.	C	k.A.	k.A.
Freestyle LibreLink – BE (Abbott SA/NV) [44]	✓	2b	2b	✓	k.A.	C	KMEHR, HL7 CDA, HL7 FHIR, SNOMED-CT	✓
Guardian Connect App (Medtronic Belgium) [45]	✓	2A	2A	✗ ⁴	k.A.	C	k.A.	k.A.
Heartfish [41]	k.A.	k.A.	k.A.	✓	k.A.	C	k.A.	k.A.
HerzMobil [39]	k.A.	k.A.	k.A.	✓	k.A.	C	HL7 CDA	✓
HumanITcare [64]	k.A.	k.A.	k.A.	✓	k.A.	C	k.A.	k.A.
icompanion (icomatrix) [65]	✓	1	k.A.	✗ ⁴	k.A.	C	k.A.	k.A.
Invirto (Sympatient GmbH) [51]	✓	1	k.A.	✓	k.A.	C	k.A.	k.A.
Kalmeda (mynoise GmbH) [52]	✓	1	k.A.	✓	k.A.	C	k.A.	k.A.

⁴ Datenschutzerklärung gemäß Datenschutz-Grundverordnung vorhanden, Speicherort der Daten außerhalb der EU.

Name der App (Hersteller)	CE Kennzeichnung*	Risikoklasse laut Hersteller*	Risikoklasse nach 2017/745 MDR*	Datenschutz**	Algorithmen***	Evidenzstufe (NICE) [11]	Interoperabilität	ELGA-kompatible Technologie
Mika (Fosanis GmbH) [66]	✓	k.A.	k.A.	✓	k.A.	C	HL7 FHIR	Nein
Mindable: Panikstörungen und Agoraphobie (Mindable Health GmbH) [53]	✓	1	k.A.	✓	k.A.	C	k.A.	k.A.
Mindmate [67]	k.A.	k.A.	k.A.	✓	k.A.	C	k.A.	k.A.
M-sense Migräne (Newsenselab GmbH) [68]	✓	1	k.A.	✓	k.A.	C	k.A.	k.A.
mySugr Tagebuch-App [46]	✓	2A	2A	✓	k.A.	C	KMEHR, HL7 CDA, HL7 FHIR, SNOMED-CT	✓
Noona (Varian Medical Systems Belgium) [69]	✓	1	k.A.	✓	k.A.	C	KMEHR, HL7 CDA, HL7 FHIR, SNOMED-CT	✓
Plateforme Maela (Medtronic Belgium) [70]	✓	1	1	✓	k.A.	C	KMEHR, HL7 CDA, HL7 FHIR, SNOMED-CT	✓
Rehappy (Rehappy GmbH) [42]	✓	1	k.A.	✓	k.A.	C	k.A.	k.A.
MobECG (Savvy) [43]	✓	2A	k.A.	k.A.	k.A.	C	k.A.	k.A.
Selfapys Online-Kurs bei Depressionen (Selfapy GmbH) [54]	✓	k.A.	k.A.	✓	k.A.	C	k.A.	k.A.
somnio (mementor DE GmbH) [55]	✓	1	k.A.	✓	k.A.	C	k.A.	k.A.
Vaica [71]	k.A.	k.A.	k.A.	✓	k.A.	C	k.A.	k.A.
velibra (GAIA AG) [56]	✓	1	k.A.	✓	k.A.	C	HL7 FHIR	Nein
Vitadio [47]	✓	k.A.	k.A.	✓	k.A.	C	k.A.	k.A.
Vivira (Vivira Health Lab GmbH) [72]	✓	1	k.A.	✓	k.A.	C	k.A.	k.A.
vorvida (GAIA AG) [57]	✓	k.A.	k.A.	✓	k.A.	C	HL7 FHIR	Nein
zanadio (aidhere GmbH) [48]	✓	1	k.A.	✓	k.A.	C	HL7 FHIR	Nein

* Laut Hersteller-Website

** Datenschutz-Statement (Einhaltung der Datenschutzgrundverordnung, inkl. Datenspeicherung in EU)

*** adaptiv, fixiert, keine Angabe

Abkürzungen: CE: europäische Konformität („conformité européenne“); GDPR: Allgemeine Datenschutz-Verordnung („General Data Protection Regulation“) k.A.: keine Angabe;

MDR: Medizinprodukteverordnung („Medical Device Regulation“)

3.3 Systematische Übersichtsarbeit zu Symptom-Checkern

In folgendem Kapitel wird zunächst die Evidenz zu DiGAs aus der Gruppe der Symptom-Checker beschrieben. Anschließend folgt im Kapitel 3.4 die Pilotevaluierung des Prozesses anhand eines ausgewählten Symptom-Checkers („Symptoma“).

**SR zu Symptom-Checkern;
Pilotevaluierung des
Prozesses**

3.3.1 Charakteristika der eingeschlossenen Studien

Basierend auf den vorab gewählten Einschlusskriterien wurde eine systematische Übersichtsarbeit zu Symptom-Checkern des britischen National Institute for Health Research (NIHR), 2019 publiziert, identifiziert [9]. In dieser systematischen Übersichtsarbeit waren 27 Studien (2 RCTs und 25 Beobachtungsstudien) inkludiert. Durch die Update-Suche wurden acht weitere Beobachtungsstudien [73-80] inkludiert.

**Update eines NIHR Berichts
von 2019: 27 Studien**

**Update: 8 weitere
Beobachtungsstudien
zum Nutzen**

**6 Studien zur
Patient*innen-Perspektive**

Zur Beurteilung der Patient*innensicht wurden außerdem weitere sechs qualitative Studien eingeschlossen [81-86]. In diesen Studien wurden die generelle Erfahrung mit Verwendung von Symptom-Checkern, positive und negative Aspekte sowie Wegbereiter und Barrieren für Symptom-Checker aus Sicht der Patient*innen erfasst.

**große Heterogenität
der Studien bzgl.
Population, Funktion,
Studiendesign, Endpunkte**

Die eingeschlossenen Studien wiesen eine große Heterogenität auf und unterschieden sich hinsichtlich der Population, der Funktionalität und Zielsetzung der untersuchten Symptom-Checker, der Anzahl der untersuchten Symptom-Checker, des Studiensettings, der Art von Komparatoren sowie der gewählten Endpunkte. Im Folgenden werden die Studiencharakteristika der eingeschlossenen Studien näher beschrieben.

Population

Die in der systematischen Übersicht 2019 [9] analysierten Symptom-Checker wandten sich an symptomatische Individuen („potentielle Patient*innen“). Sieben der zusätzlich inkludierten Studien [73-75, 77-80] analysierten Symptom-Checker, die von symptomatischen Individuen („potentiellen Patient*innen“) zu verwenden sind. Eine der inkludierten Studien analysierte aber einen Symptom-Checker, der von Gesundheitspersonal verwendet wurde [76].

**symptomatische
Individuen („potentielle
Patient*innen“),
aber auch
Gesundheitspersonal**

Funktionalität und Zielsetzung

Die in den Studien untersuchten Symptom-Checker unterscheiden sich in ihrer Funktionalität und Zielsetzung und lassen sich dabei in drei Gruppen einteilen:

3 Gruppen von Funktionen:

- Symptom-Checker, welche Diagnosevorschläge erstellen, ohne eine Triage-Empfehlung anzugeben,
- Symptom-Checker, welche eine Triage-Empfehlung erstellen, ohne dabei Diagnosevorschläge anzugeben,
- Symptom-Checker, welche sowohl Diagnosevorschläge als auch eine Triage-Empfehlung erstellen.

*** Diagnosevorschläge
ohne Triage-Empfehlung
* Diagnosevorschläge
mit Triage-Empfehlung
* Triage-Empfehlung
ohne Diagnosevorschläge**

entsprechend der
Funktion: Endpunkte

Dementsprechend wurden von den Studien – je nach Funktionalität der untersuchten Symptom-Checker – als untersuchte Endpunkte die diagnostische Genauigkeit der Diagnosevorschläge und/oder die Angemessenheit und Sicherheit der Triage-Empfehlungen erhoben. Eine detaillierte Beschreibung der Endpunkte der eingeschlossenen Studien findet sich im Methodenkapitel 2.3.4.

Art und Anzahl der untersuchten Symptom-Checker

NIHR-Bericht:

In der systematischen Übersicht des NIHR [9] wurden 27 Studien (29 Publikationen) in der systematischen Suche eingeschlossen:

17 Studien zu einzelnen
Symptom Checkern

- 17 Studien untersuchten Symptom-Checker als eigenständige Intervention (davon neun generische und acht diagnose-spezifische Symptom-Checker).

5 Studien als Add-On zu
klinischer Beurteilung

- Fünf Studien untersuchten Symptom-Checker als Baustein einer erweiterten klinischen Beurteilung (im Rahmen einer elektronischen Konsultation), um zu beurteilen, ob die Symptome der Patient*innen einen Ärzt*innenbesuch erfordern oder eine Selbstbehandlung erfolgen kann.

5 Studien zur
Patient*innen-Perspektive

- Fünf Studien untersuchten die Patient*innenperspektive zum Thema Symptom-Checker und Selbstdiagnosen.

Update Studien:

5 Studien zu einzelnen
Symptom Checkern;

3 Studien zu Vergleichen
zwischen
unterschiedlichen
Symptom Checkern

Die Anzahl der untersuchten Symptom-Checker in den acht zusätzlich inkludierten Studien betrug folgendermaßen: Fünf der inkludierten Studien [73, 74, 76-78] untersuchten konkret jeweils einen Symptom-Checker. Diese untersuchten Symptom-Checker waren „Babylon Triage and Diagnostic System“ [73], „Symptoma“ [74], „Ada DX“ [76], „Algorithm-generated self-triage-score“ [77] und „Should I see a doctor?“ [78]. Drei der inkludierten Studien [75, 79, 80] zogen hingegen einen Vergleich von mehreren Symptom-Checkern. In diesen Studien wurden acht Symptom-Checker [75], 36 Symptom-Checker [79] sowie 15 Symptom-Checker [80] miteinander verglichen.

Studiendesigns, -setting und Komparatoren

NIHR-Bericht:
2 RCTs +

25 Beobachtungsstudien

In der systematischen Übersichtsarbeit des NIHR [9] wurden zwei RCTs und 25 Beobachtungsstudien inkludiert. Die diagnostische Genauigkeit wurde teilweise in Simulationen (unter Verwendung von klinischen Fallvignetten), teilweise im direkten Vergleich mit Ärzt*innen erhoben; dabei wurden die Diagnosevorschläge und das empfohlene Triage-Level in diesen Fallvignetten-Studien erhoben.

Fallvignetten-Studien

Web-Applikationen,
SmartPhone App

Die meisten der untersuchten Symptom-Checker waren durch Websites aufrufbar, ein Symptom-Checker war durch Smartphone-App verfügbar. Einige der in den inkludierten Studien untersuchten Symptom-Checker waren nicht mehr verfügbar. Zur genauen Funktionsweise der Symptom-Checker wurden in den inkludierten Studien kaum Angaben gemacht.

8 zusätzliche Primärstudien:

5 Fallvignetten-Studien;
2 prospektive

Beobachtungsstudien;
1 Fall-Kontroll-Studie

Die acht zusätzlich inkludierten Primärstudien wiesen eine Vielzahl an Studiendesigns auf. Das Vorliegen von klinischen Fall-Vignetten zur Messung von diagnostischer Genauigkeit in Rollenspielszenarien erschwerte eine eindeutige Klassifizierung in bereits etablierte Studientypen: nachfolgend werden diese Studien im Studientyp als Fallvignetten-Studien beschrieben.

Fünf der inkludierten Studien verwendeten im Studiensetting klinische Fallvignetten [73-75, 79, 80]. Klinische Fallvignetten sind fiktive Beispiele von Patient*innenfällen mit Symptomen, zugehöriger Diagnose und ggf. zuge-

höriger Triage-Empfehlung. Klinische Fallvignetten können etwa bei Prüfungen in der Ausbildung von Mediziner*innen zum Einsatz kommen, oder um Differenzen in Diagnose- und Triage-Empfehlungen zwischen Arztpraxen zu erheben [87].

Bei den inkludierten Studien mit klinischen Fallvignetten gab es Unterschiede in der Anzahl der verwendeten Fallvignetten und der Art des Komparators. In den Beobachtungsstudien wurden 200 klinische Fallvignetten [75], 100 klinische Fallvignetten [73], 48 klinische Fallvignetten [79] und in zwei Studien 45 klinische Fallvignetten [74, 80] verwendet.

Die Robustheit der Studiendesigns mit fiktiven, klinischen Fallvignetten ist einerseits durch die Rollenspielszenarien, andererseits durch heterogene Komparatoren eingeschränkt. Eine Analyse des Verzerrungspotenzials der Studien findet sich im Anhang Tabelle Tabelle A-5. In den Studien über Symptom-Checker, welche im Studiendesign klinische Fallvignetten verwendeten, wurde als Komparator entweder direkt mit der gesuchten Diagnose der Fallvignette verglichen oder die Einschätzung von Gesundheitspersonal oder Laien erhoben:

- In einer Beobachtungsstudie [73] wurden in einem Rollenspielszenario mit sieben Allgemeinmediziner*innen und zehn Patient*innen (gespielt durch Allgemeinmediziner*innen) die klinischen Fallvignetten mit Diagnosevorschlägen versehen und mit Diagnosevorschläge des Symptom-Checkers verglichen.
- In einer weiteren Beobachtungsstudie [74] wurden als Komparator für die klinischen Fallvignetten andere Symptom-Checker gewählt (die Studie wurde als Aktualisierung einer Studie aufgesetzt, in welcher 19 Symptom-Checker anhand 45 klinischen Fallvignetten zur diagnostischen Genauigkeit getestet wurden [14]).
- In einer weiteren Beobachtungsstudie [75] wurde die Einschätzung zu den klinischen Fallvignetten von acht Symptom-Checkern mit der Einschätzung von sieben Allgemeinmediziner*innen verglichen.
- In einer weiteren Beobachtungsstudie [79] wurde die Einschätzung zu klinische Fallvignetten von 36 Symptom-Checkern erhoben, ohne dabei mit einer Einschätzung von Gesundheitspersonal zu den klinischen Fallvignetten zu vergleichen.
- In einer weiteren Beobachtungsstudie [80] wurde die Einschätzung zu klinischen Fallvignetten von 15 Symptom-Checkern mit der Einschätzung von 91 Laien (Laien ohne professionellen medizinischen Hintergrund) als Komparator verglichen.

Die übrigen drei der inkludierten Studien verwendeten in der Erhebung keine klinischen Fallvignetten:

- Eine Beobachtungsstudie [77] schloss 492 Patient*innen, die ohne Krankenwagen in einer Notaufnahme vorstellig wurden, ein. Diesen Patient*innen wurde anhand des „Algorithmus-generated self-triage score“ Fragen gestellt, die einen Triage-score generierten. Als Komparator diente die Einschätzung von Gesundheitspersonal („Triage nurses“) unter Verwendung der kanadischen „Triage- und Dringlichkeitsskala“.
- Eine Beobachtungsstudie [78] erhob mittels integriertem Fragebogen in der App des Symptom Checkers, ob die Nutzer*innen beabsichtigen, die Triage-Empfehlung des Symptom-Checkers zu befolgen. 4.456 Nutzer*innen nahmen mittels des integrierten Fragebogens an der

**Anzahl der Fallvignetten:
45 bis 200 Fälle**

**Studiendesigns
und Komparatoren:**

**Fallvignetten mit
Ärzt*innen,
Gesundheitspersonal
oder auch Laien
als Vergleich**

**Patient*innen in
Notaufnahme
(Selbsteinweisung) mit
Einschätzung von
Gesundheitspersonal
verglichen**

**oder
Vergleich mit
Einschätzung von
Triage-Telefondienst**

oder

Vergleich mit
Gesundheitsakten aus
Ambulanz für seltene
Erkrankungen

Erhebung teil. Als Komparator wurden 126 der 4,456 Teilnehmer*innen von Gesundheitspersonal („Triage nurses“) angerufen, um zu bewerten, ob die Empfehlungen des Symptom-Checkers mit der Empfehlung des Triage-Anrufes übereinstimmen.

- Eine Beobachtungsstudie [76] wählte aus elektronischen Gesundheitsakten 93 Fälle mit Diagnosen von seltenen Erkrankungen einer Ambulanz für seltene inflammatorische Systemerkrankungen ein. Die Auswahl der Fälle erfolgte subjektiv mit Priorisierung von Fällen mit diagnostischen Odysseen, d. h. Patient*innenfälle, bei welchen erst nach längerer Zeit die korrekte Diagnose gefunden wurde. Ein Symptom-Checker machte anhand der dokumentierten Symptome in den elektronischen Gesundheitsakten Diagnosevorschläge. Die Genauigkeit der Diagnosevorschläge wurde zum dokumentierten Zeitpunkt der Diagnose aus der elektronische Gesundheitsakte verglichen (Zeit bis zur bestätigten Diagnose ohne Einsatz des Symptom-Checkers als Intervention).

Qualität der eingeschlossenen Studien

NIHR-Bericht:
hohe Qualität

Die systematische Übersichtsarbeit des NIHR-Berichts [9] wurde mit dem AMSTAR-II Assessment Tool [27] mit hoher Qualität bewertet (vgl. Appendix „Qualitätsbewertung des NIHR Berichts ...“).

Update Studien:
7 Studien hohes RoB
1 Studie niedriges RoB

In den zusätzlich identifizierten Beobachtungsstudien wurden sieben Beobachtungsstudien [73-77, 79, 80] mit hohem Risk-of-Bias (RoB) und eine Beobachtungsstudie [78] mit niedrigem RoB bewertet (vgl. Tabelle A-5).

Endpunkte der eingeschlossenen Studien

Für die eingeschlossenen Studien wurden die Wirksamkeit, die Sicherheit, die diagnostische Genauigkeit, die Genauigkeit der Triage-Empfehlungen, systemisch/organisatorische Endpunkte sowie patient*innenrelevante Endpunkte als relevant erachtet. Eine detaillierte Beschreibung dieser Endpunkte findet sich im Methodenkapitel 2.3.4.

Studiencharakteristika

Die wesentlichen Charakteristika der systematischen Übersichtsarbeit des NIHR-Berichts [9] und der zusätzlich identifizierten Studien sind im Überblick in den nachfolgenden Tabellen (Tabelle 3-8 und Tabelle 3-9) dargestellt.

Tabelle 3-8: Wesentliche Merkmale der systematischen Übersichtsarbeit des NIHR-Berichts

Autor*innen, Jahr,	Chambers et al., 2019 [9]
Studiendesign	Systematic Review
Qualität (AMSTAR-II)	Hoch
Intervention	Symptom-Checker
Komparator	NHS 111 (telefonische Gesundheitsberatung), keine Restriktion
Setting	Allgemeinbevölkerung mit „dringendem“ Gesundheitsproblem (akut und chronisch)
Suchzeitraum	Bis 2018
Eingeschlossene Studien, Qualität der Evidenz	27, niedrig (2 RCTs, 25 Beobachtungsstudien)

Abkürzungen: NHS – National Health Service; RCT – randomisiert kontrollierte Studie („randomized controlled trial“)

Tabelle 3-9: Wesentliche Merkmale des Studientyps, Studiensettings, verwendeten Interventionen und gemessenen Endpunkten der zusätzlich identifizierten Studien

Autor*innen, Jahr	Studientyp	Studiensetting	Intervention	Endpunkte	Risk-of-Bias (QUADAS-II)
Schmieding et al., 2021 [80]	Prospektive Fallvignetten-Studie ⁵	Vergleich der Bewertung von klinischen Fallvignetten durch Symptom-Checker und durch Laien ohne professionellen medizinischen Hintergrund.	15 Symptom-Checker	„Genauigkeit“ der Triage-Empfehlungen	Hoch
Trivedi et al., 2021 [77]	Prospektive Beobachtungs-Studie	Ermittlung eines Triage-Scores für Patient*innen, die in einer Notaufnahme vorstellig wurden. Triage-Score des Symptom-Checkers verglichen mit der Triage-Einschätzung des Gesundheitspersonals.	Algorithmischer Fragebogen mit Algorithmus-generiertem Self-Triage-Score	Genauigkeit der Triage-Empfehlungen	Hoch
Baker et al., 2020 [73]	Prospektive Fallvignetten-Studie ⁵	Scheinkonsultationen (Rollenspielszenario) mit klinischen Fallvignetten. Beurteilung von klinischen Fallvignetten durch Ärzt*in verglichen mit Beurteilung durch Symptom-Checker.	Babylon Triage and Diagnostic System	„Genauigkeit“ der Diagnosevorschläge „Genauigkeit“ der Triage-Empfehlungen	Hoch
Gilbert et al., 2020 [75]	Prospektive Fallvignetten-Studie ⁵	Beurteilung von klinischen Fallvignetten durch Ärzt*in verglichen mit Beurteilung durch Symptom-Checker.	8 Symptom-Checker	„Genauigkeit“ der Diagnosevorschläge „Genauigkeit“ der Triage-Empfehlungen	Hoch
Hill et al., 2020 [79]	Retrospektive Fallvignetten-Studie ⁵	Beurteilung von klinischen Fallvignetten durch Symptom-Checker.	36 Symptom-Checker	„Genauigkeit“ der Diagnosevorschläge „Genauigkeit“ der Triage-Empfehlungen	Hoch
Nateqi et al., 2019 [74]	Retrospektive Fallvignetten-Studie ⁵	Aktualisierung einer Studie, in welcher 19 Symptom-Checker klinische Fallvignetten beurteilten (Ergänzung durch fünf weitere Symptom-Checker).	24 Symptom-Checker	„Genauigkeit“ der Diagnosevorschläge	Hoch
Ronicke et al., 2019 [76]	Fall-Kontroll-Studie	Fälle aus elektronischen Gesundheitsakten einer Ambulanz für seltene inflammatorische Systemerkrankungen wurden mit dem Symptom-Checker beurteilt. Die Zeit bis zur Diagnose wurde berechnet als der Zeitraum zwischen dem ersten dokumentierten Besuch mit Symptomen und dem dokumentierten Besuch der Diagnose.	„Ada DX“	Genauigkeit der Diagnosevorschläge (zum Zeitpunkt der Dokumentation der klinischen Diagnose)	Hoch
Verzantvoort et al., 2018 [78]	Prospektive Cross-Sectional Studie	Nutzer*innen des Symptom-Checkers wurden von Gesundheitspersonal angerufen; die Triage-Empfehlungen der App wurden mit den Empfehlungen der Triage-Pflege*in verglichen.	Smartphone App „Should I see a doctor?“ (in Dutch: „moet ik naar de dokter?“)	Genauigkeit der Triage-Empfehlungen	Niedrig

⁵ Bei den inkludierten Fallvignetten-Studien handelt es sich um Audit-Studien, in welchen ein Abgleich mit klinischen Fallvignetten zur Erhebung der Übereinstimmung mit Diagnosevorschlägen und Triage-Empfehlungen durchgeführt wird.

3.3.2 Evidenz der eingeschlossenen Studien

**zusammenfassende
Evidenz aus NIHR-Bericht
und den zusätzlichen
Studien**

Aufgrund der Heterogenität der inkludierten Studien wurde in der systematischen Übersichtsarbeit des NIHR-Berichts [9] von einer Meta-Analyse abgesehen und eine narrative Evidenzsynthese durchgeführt. Eine Zusammenfassung dieser Evidenz des NIHR-Berichts findet sich in der Datenextraktionstabelle im Anhang in Tabelle A-2. Im folgenden Abschnitt soll diese Evidenzsynthese durch die acht zusätzlichen Primärstudien [73-80] und den sechs zusätzlich identifizierten Studien zur Patient*innensicht für die vorab definierten Endpunkte ergänzt werden. Die Datenextraktionstabellen zu den zusätzlich identifizierten Studien finden sich im Anhang (siehe Tabelle A-3 und Tabelle A-4).

Wirksamkeit

**1 RCT:
geringfügig längere
Krankheitsdauer mit
Symptom-Checker,
n.s. Unterschied**

Der Endpunkt klinische Wirksamkeit (Krankheitsdauer oder Schwere der Erkrankung) wurde in der systematischen Übersichtsarbeit des NIHR [9] von einem RCT berichtet. In diesem war die Krankheitsdauer in der Interventionsgruppe mit Symptom-Checker bei 11,3 Tagen vs. 10,9 Tagen in der Kontrollgruppe. Dieses Ergebnis wurde als statistisch nicht signifikant angegeben. In den acht zusätzlich identifizierten Primärstudien [73-80] wurde der Endpunkt klinische Wirksamkeit nicht erhoben.

Sicherheit

**in 6 Studien erhoben:
keine Aussagen möglich,
u. a. wegen Mängel in
Studien**

Der Endpunkt Sicherheit (Schäden durch falsch-positive/falsch-negative Diagnosevorschläge oder Triage-Empfehlungen) wurde in der systematischen Übersichtsarbeit des NIHR [9] von sechs inkludierten Beobachtungsstudien berichtet. In diesen Studien konnte keine Evidenz zu möglichen Schäden durch Verwendung von Symptom-Checkern identifiziert werden. Etwaige Schäden durch Symptom-Checker waren jedoch nicht auszuschließen, da die inkludierten Studien Limitationen aufwiesen (Studienpopulation gesunde Student*innen mit einer beschränkten Auswahl an Diagnosen oder simulierte Studiensettings mit Fallvignetten). In den acht zusätzlich identifizierten Beobachtungsstudien [73-80] wurde der Endpunkt Sicherheit nicht erhoben.

Diagnostische Genauigkeit

**NIHR-Bericht:
niedrige diagnostische
Genauigkeit

34 % Übereinstimmung
bei Top1,
58 % bei Top20
diagnostischen
Vorschlägen der
Symptom Checker**

Der Endpunkt diagnostische Genauigkeit wurde in der systematischen Übersichtsarbeit des NIHR [9] von sieben inkludierten Beobachtungsstudien als Übereinstimmung mit der klinischer Diagnose definiert. Die diagnostische Genauigkeit wurde bei der Evaluation von klinischen Fallvignetten, sowie im Vergleich mit dem Gesundheitspersonal, als niedrig angegeben. Die umfassendste Bewertung einer Beobachtungsstudie [14] untersuchte 23 Symptom-Checker über 770 standardisierte Patient*innenfälle und fand eine diagnostische Genauigkeit mit der korrekten Diagnose als Top 1 Vorschlag in 34 % der Fälle (95 % CI: 31 %-37 %), innerhalb der Top 20 Vorschläge in 58 % der Fälle (95 % CI: 55 %-62 %). Im direkten Vergleich mit Ärzt*innen, unter Verwendung von 43 klinischen Fallvignetten, konnten die Ärzt*innen die korrekte Diagnose besser als die Symptom-Checker erfassen: im Top 1 Vorschlag 72,1 % vs. 34 % ($p < 0,001$), im Top 3 Vorschlag 84,3 % vs. 51,2 % ($p < 0,001$). Eine Ausnahme stellte eine Erhebung in einem Gesundheitszentrum für Studierende dar: hier wurde für 59 Studierende eine hohe Übereinstimmung zwischen dem „24/7 WebMed system“ und der Einschätzung durch Gesundheitspersonal

gefunden (Kappa Werte 0,89-0,94); als mögliche Begründung für die hohe Übereinstimmung wird die kleine, relativ gesunde Studienpopulation angegeben.

Die Sensitivität (Anteil an tatsächlich Kranken) der Genauigkeit der Diagnosevorschläge wurde in zwei Studien⁶ [73, 75] berichtet. In der ersten Studie [73] lag der Durchschnitt der Ärzt*innen bei 83,9 % (95 % CI: 75,6 %-92,3 %), die Erkennung durch den Symptom-Checker „Babylon AI“ bei 80,0 %. In der zweiten Studie [75] lag die Sensitivität der Diagnosevorschläge: bei 23,5 %- 99,0 %. Die Spezifität (Anteil gesunder Personen, die auch ein negatives Testergebnis erhält) der Genauigkeit der Diagnosevorschläge wurde in keiner Studie berichtet.

Der positiv prädiktive Wert der diagnostischen Genauigkeit wurde in einer Studie [73] berichtet. In dieser Studie lag der Durchschnitt der Ärzt*innen bei 43,6 % (95 % CI: 36,3 %-50,9 %), die Erkennung durch den Symptom Checker „Babylon AI“ bei 44,4 %. Der negativ prädiktive Wert der diagnostischen Genauigkeit wurde in keiner der inkludierten Studien berichtet.

In den acht zusätzlich identifizierten Beobachtungsstudien [73-80] wurde der Endpunkt Übereinstimmung mit klinischer Diagnose in vier Beobachtungsstudien [74, 75, 79] gemessen.

- In der ersten Studie [74] wurde die Genauigkeit der Diagnosevorschläge in den Top 1, Top 3 und Top 10 der Vorschläge der Symptom Checker erhoben, diese lagen für den Top 1 Vorschlag zwischen 22 %-82,2 %, für die Top 3 Vorschläge zwischen 31 %-100 % und für die Top 10 Vorschläge zwischen 38 %-100 %.
- In der zweiten Studie [75] wurde die Genauigkeit der Diagnosevorschläge in den Top 3 erhoben. Der Durchschnitt der Ärzt*innen lag hier bei 82,1 % ± 5,2 %. Für die Symptom Checker lag die Genauigkeit zwischen 23,5 %-70,5 %.
- In der dritten Studie [79] wurden die Ergebnisse als Durchschnitt der 27 untersuchten diagnostischen Symptom-Checker für die Top 1, Top 3 und Top 10 Treffergenauigkeit angegeben. Dabei lag die korrekte Diagnose an 1. Stelle in 36 % (95 % CI: 31 %-42 %), innerhalb Top 3 in 52 % (95 % CI: 47 %-59 %) und innerhalb Top 10 in 58 % (95 % CI: 53 %-65 %). Außerdem wurden die Symptom-Checker in dieser Studie in Symptom-Checker mit oder ohne künstliche Intelligenz eingeteilt, dabei zeigte sich ein Unterschied: Korrekte Diagnose an 1. Stelle in 46 % (95 % CI: 40 %-57 %) der Fallvignetten durch Symptom-Checker mit künstlicher Intelligenz. Im Gegensatz dazu die korrekte Diagnose an 1. Stelle in 32 % (95 % CI: 26 %-38 %) der Fallvignetten durch Symptom-Checker ohne künstliche Intelligenz.
- Die Übereinstimmung der Diagnosevorschläge zum Zeitpunkt der klinischen Diagnose wurde in der vierten Studie [76] erhoben: Noch vor dem Zeitpunkt der dokumentierten Diagnose in der elektronischen Gesundheitsakte kam es durch den Symptom Checker „Ada“ als Top 1 Diagnosevorschlag in 37,6 % der Fälle, unter den Top 5 Vorschlägen in 53,8 %. Zum Zeitpunkt des ersten dokumentierten Patient*innenbesuchs (laut elektronischer Gesundheitsakte) kam es durch den Symptom Checker als Top 1 Diagnosevorschlag in 16,1 %, unter den Top 5 Vorschlägen in 33,3 % der Fälle. Als medianer Vorteil der kor-

Sensitivität:
23,5 % bis 99 %

Spezifität:
nicht berichtet

positiv prädiktiver Wert:
gleich wie bei Ärzt*innen
negativ prädiktiver Wert:
nicht berichtet

Update
4 Studien zu
diagnostischer Genauigkeit

große Unterschiede
zwischen einzelnen
Symptom Checkern

tendenziell sehr geringe
Übereinstimmung

jedenfalls geringer
als durch Ärzt*innen
(wenn verglichen)

bessere Ergebnisse bei
Symptom Checkern mit
künstlicher Intelligenz

⁶ Eine weitere Studie [74] berichtete die Sensitivität und Spezifität als Abbildung dargestellt.

rekten Diagnosevorschläge im Vergleich zum Zeitpunkt der klinischen Diagnose wurde von den Autor*innen ein Zeitraum von 3 Monaten (50 %) für Top 5 Vorschläge, und ein Zeitraum von 1 Monat (21 %) für den Top 1 Diagnosevorschlag angegeben.

Triage-Empfehlungen

<p>NIHR-Bericht: Symptom Checker „übersichtlich“ im Vergleich zu Triage</p> <p>Entscheidung von Gesundheitspersonal</p> <p>dringende Fälle werden erkannt</p>	<p>In der systematischen Übersichtsarbeit des NIHR-Berichts [9] wurden die Endpunkte betreffend Triage-Empfehlungen von sechs inkludierten Beobachtungsstudien berichtet. Zum Großteil wurden die Triage-Einschätzungen der Symptom-Checker im Vergleich zu Gesundheitspersonal als unterlegen und übersichtlich eingestuft. Die umfassende Bewertung von 23 Symptom-Checker [14] fand eine angemessene Triage-Empfehlung in 57 % (95 % CI: 52 %-61 %) der Fälle. In einer weiteren Evaluation von vier Symptom-Checkern wurde eine Range der Triage-Übereinstimmung mit Gesundheitspersonal von 30 %-95 % angegeben. Eine Evaluation von drei Symptom-Checkern fand eine Triage-Übereinstimmung bei abdominalen Schmerzen in 33 % der Fälle, in 39 % eine dringendere Empfehlung und in 30 % eine weniger dringende Empfehlung als Gesundheitspersonal. Eine weitere Evaluation fand mit 39 % eine Übereinstimmung der Triage-Empfehlungen der Symptom-Checker mit dem Gesundheitspersonal. Ein weiterer Symptom-Checker empfiehlt in 85 % das Aufsuchen eine*r Ärzt*in, auch wenn nach Einschätzung von Gesundheitspersonal eine Selbstbehandlung ausgereicht hätte. Die einzigen Studien, die eine gleichwertige oder überlegene Triage-Empfehlung berichteten, wurden vom Hersteller selbst durchgeführt: Ein Symptom-Checker gab bei der Analyse von klinischen Fallvignetten die korrekte Einschätzung in 88,2 % der Fälle (verglichen mit 75,5 % von Ärzt*innen und 73,5 % von Pflegepersonal) an. Laut einer darauf aufbauenden Analyse des NHS England waren alle Triage-Empfehlungen des Symptom-Checkers für eine dringende oder für eine Notfall-Behandlung angemessen.</p>
<p>Update</p> <p>6 Studien: Angemessenheit und Sicherheit wenig einheitlich definiert</p> <p>übersichtlich = sicher?</p>	<p>In den acht zusätzlich identifizierten Beobachtungsstudien [73-80] wurden die Endpunkte betreffend Triage-Empfehlungen in sechs Studien [73, 75, 77-80] berichtet. In diesen Beobachtungsstudien wurden die Sicherheit und die Angemessenheit der Triage-Empfehlungen erhoben. Diese Endpunkte wurden von den inkludierten Studien jedoch nicht einheitlich definiert. So wurde in einer Beobachtungsstudie [73] die Triage-Sicherheit danach beurteilt, ob die vorgeschlagene Triage-Aktion als sicher angesehen wurde, auch wenn sie übersichtlich war. Für eine andere Beobachtungsstudie [75] hingegen war die Triage-Sicherheit dann gegeben, wenn die Empfehlung des Symptom-Checkers maximal eine Stufe⁷ weniger konservativ als die Gold-Standard-Vignettenempfehlung war.</p> <ul style="list-style-type: none"> ■ Die Sensitivität, Spezifität, positiv prädiktiver Wert und negativ prädiktiver Wert zu Triage-Empfehlungen wurde von einer Studie [78] angegeben. In dieser Studie wurde die App „Should I see a doctor?“ untersucht. Die Sensitivität lag bei 84 % (95 % CI: 74 %-91 %), die Spezifität bei 74 % (95 % CI: 58 %-86 %), der positiv prädiktive Wert bei 88 % (95 % CI: 79 %-94 %) und der negativ prädiktive Wert bei 67 % (95 % CI: 51 %-81 %).
<p>Sensitivität: 84 %</p> <p>Spezifität: 74 %</p> <p>positiv prädiktiver Wert: 88 %</p> <p>negativ prädiktiver Wert: 67 %</p>	

⁷ Verwendete Triage-Stufen: 1 = Notruf setzen; 2 = in Notfallambulanz begeben; 3 = Allgemeinmedizin innerhalb von vier Stunden; 4 = Allgemeinmedizin noch am selben Tag; 5 = Allgemeinmedizin ohne Dringlichkeit; 6 = Selbstversorgung. Triage-Sicherheit wurde definiert als maximal eine Stufe weniger konservativ als die Gold-Standard-Vignettenempfehlung.

- In einer prospektiven Fallvignetten-Studie [73] lag der Durchschnitt der Übereinstimmung der Triage-Empfehlungen einer unabhängigen Person für Ärzt*innen bei 93,1 % (95 % CI: 89,6 %-96,6 %), für den Symptom-Checker bei 97,0 %.
- In einer weiteren prospektiven Fallvignette-Studie [75] lag der Durchschnitt der Übereinstimmung der Triage-Empfehlungen für Ärzt*innen bei 97,0 % ± 2,5 %, für die untersuchten Symptom-Checker bei 90,1 % ± 7,4 %.
- Für die prospektive Fallvignetten-Studie welche den Algorithmus-generierten Self-Triage-Score untersuchte [77], gab es in 27 % der Fälle eine Übereinstimmung der Triage-Empfehlungen. In 65 % der Fälle wurde die Dringlichkeit nach Verwendung des Symptom-Checkers überschätzt. In 8 % der Fälle wurde die Dringlichkeit nach Verwendung des Symptom-Checkers unterschätzt. 41 % der Patient*innen schätzten nach Verwendung des Symptom-Checkers, dass eine stationäre Aufnahme für sie indiziert sei, hingegen kam es durch Evaluation der Symptome durch Gesundheitspersonal in 17 % der Fällen zu einer tatsächlichen Aufnahme.
- In einer prospektiven Cross-Sectional-Studie [78] entsprach die Triage-Empfehlung des Symptom-Checkers in 81 % der Fällen der Empfehlung des Triage-Anrufs. In 8 % der Fälle wurde die Dringlichkeit nach Verwendung des Symptom-Checkers überschätzt, in 11 % unterschätzt.
- In der retrospektiven Fallvignetten-Studie [79] kam es in 49 % (95 % CI: 44 %-54 %) der Fälle zu einer korrekten Triage-Empfehlung. Die korrekte Empfehlung für Notfallversorgungen lag bei 63 % (95 % CI: 52 %-71 %). Die korrekte Empfehlung für dringende Versorgung lag bei 56 % (95 % CI: 52 %-75 %), für nicht dringende Versorgung bei 30 % (95 % CI: 11 %-39 %) und schließlich für Selbstversorgung bei 40 % (95 % CI: 26 %-49 %).
- Für die prospektive Fallvignetten-Studie mit dem Vergleich der Triage-Einschätzung durch Laien [80] lag die mittlere Triage-Genauigkeit durch Laien bei 60,9 % (SD 6,8 %; 95 % CI: 59,5 %-62,3 %, die der Symptom-Checkers bei 58 % (SD 12,8 %). Die Erkennung von Notfällen durch Laien lag bei 67,5 % (SD 16,4 %; 95 % CI: 64,1 %-70,8 %), die der Symptom-Checker bei 80,6 % (SD 17,9 %). Die Erkennung von falsch-positiven Notfällen durch Laien lag bei 4,10 %, die der Symptom-Checker bei 24,7 %.

Überschätzung der Dringlichkeit durch Symptom Checker

Systemisch/organisatorische Auswirkungen

Endpunkte (Anzahl Arzt-/Ambulanzbesuche) zu den systemisch/organisatorischen Auswirkungen bei Verwendung von Symptom-Checker wurden in der systematischen Übersichtsarbeit des NIHR-Berichts [9] von einem RCT und sechs Beobachtungsstudien berichtet.

- In dem RCT wurde gezeigt, dass die Interventionsgruppe (mit Symptom-Checker) weniger oft Arzt-/Ambulanzbesuche hatte, obwohl die Krankheitsdauer länger war. In der Interventionsgruppe kam es auch häufiger zu telefonischen Kontakt mit den behandelnden Ärzt*innen, nachdem der Symptom-Checker die Patient*innen dazu aufgefordert hat.
- Eine Studie argumentierte, dass der Einfluss der Symptom-Checker auf die Anzahl der Arzt-/Ambulanzbesuche nicht messbar sei.

NIHR Bericht

Tendenz:
geringere
Inanspruchnahme von
Arzt-/Ambulanzbesuchen
häufigere
Kontaktaufnahme bei
Aufforderung durch
Symptom-Checker

- häufiger Selbstbehandlung**
- keine zusätzliche Evidenz in Update**
- Eine Studie fand, dass 18 % der Nutzer*innen welche einen Arzt-/Ambulanzbesuch planteten, diesen nach Verwendung des Symptom-Checkers nicht durchführten. Zusätzlich gaben 14 % der Nutzer*innen an, dass sie ohne den Symptom-Checker eine*n Ärzt*in oder Ambulanz aufgesucht hätten.
 - Eine weitere Studie an vier Symptom-Checker im Vergleich zu telefonischer Gesundheitsberatung fand eine häufigere Empfehlung der Symptom-Checker für Selbstbehandlung (18 % vs. 14 %) und eine seltenere Empfehlung für eine allgemeine Behandlung (40 % vs. 60 %).
 - Laut der Studie eines Herstellers führten die Empfehlungen des Symptom-Checkers im Vergleich zur telefonischen Gesundheitsberatung häufiger zu Selbstbehandlung (40 % vs. 14 %).
 - In einer weiteren Studie eines Symptom-Checkers für ein Gesundheitszentrum für Studierende wurde gezeigt, dass in Fällen in denen durch den Symptom-Checker eine Selbstbehandlung empfohlen wurde, eine Reduktion der Arzt-/Ambulanzbesuche möglich sei.
 - In einer weiteren Studie konnte gezeigt werden, dass Studierende, welche eine Selbst-Diagnose über einen Symptom-Checker oder eine Google-Suche erhielten, ein größeres Bedürfnis hatten die Diagnose medizinisch abzuklären.

In den acht zusätzlich inkludierten Primärstudien [73-80] wurden keine systemisch/organisatorischen Endpunkte berichtet.

Patient*innensicht

- NIHR Bericht:**
- Patient*innen befolgen Empfehlung des Symptom-Checkers**
- In der systematischen Übersichtsarbeit des NIHR [9] wurde in zwei inkludierten Studien die Patient*innensicht beschrieben. Dabei wurde berichtet, ob die Patient*innen beabsichtigen, die von den Symptom-Checkern vorgeschlagenen Triage-Empfehlungen zu befolgen. In einer Studie wurden Ergebnisse aus qualitativen Studien berichtet, die darauf hindeuten, dass Nutzer*innen generell die Triage-Empfehlungen der Symptom-Checker befolgen würden. Bei Empfehlung des Symptom-Checkers für eine dringende oder Notfall-Behandlung gaben die Nutzer*innen an, sich zunächst an Allgemeinmediziner*innen zu wenden. In der zweiten Studie gaben 57 % der Patient*innen an, der Triage-Empfehlung des Symptom-Checkers gefolgt zu sein.
- Anwender-freundlichkeit: grundsätzlich Zufriedenheit**
- Eine der zusätzlich identifizierten Studien [78] berichtete die Patient*innensicht zur Anwenderfreundlichkeit der digitalen Gesundheitsanwendung. In diesem Survey wurde die App von 87 % Patient*innen als ‚neutral‘ bis ‚sehr klar‘ bewertet. 89 % der Patient*innen waren mit der App ‚neutral‘ bis ‚sehr zufrieden‘. Patient*innen, welche mit der App unzufrieden waren, gaben folgende Begründungen an: für 62 % der Patient*innen konnten die Symptome bzw. die Krankheit nicht vollständig geschildert werden, für 24 % waren die Empfehlungen zu schnell und nicht umfassend, von 14 % wurden strukturelle Mängel angegeben. In derselben Studie [78] gaben 65 % der Patient*innen an, die Empfehlungen des Symptom-Checkers befolgen zu wollen. Von den Patient*innen, welche keine Absicht hatten, die Empfehlung zu befolgen, wurden folgende Gründe angegeben: für 33 % konnte die Krankheit nicht vollständig geschildert werden, für 27 % fand bereits eine Ärzt*in-Konsultation statt, und 27 % gaben an mehr Vertrauen in das eigene Urteilsvermögen zu haben.
- wenn unzufrieden: Symptome der Krankheit zu wenig berücksichtigt**
- zu rasche Empfehlung**

In sechs weiteren Studien [81-86] wurden Patient*innen zu ihren generellen Erfahrungen mit Symptom-Checker, positive und negative Aspekte sowie empfundene Wegbereiter und Barrieren für die Verwendung von Symptom-Checker befragt. In diesen Studien wurde gezeigt, dass die meisten Nutzer*innen bisher wenig Erfahrung mit Symptom-Checkern haben. Die Nutzer*innen waren tendenziell mit der Bedienbarkeit der Benutzeroberfläche zufrieden. Das Symptom-Assessment wurde von den Nutzer*innen als strukturiert und personalisiert beschrieben; das Fehlen der verbalen Interaktion mit dem Gesundheitspersonal und die eingeschränkte Möglichkeit der vollständigen Symptom-Beschreibung wurden als Hindernisse für die Verwendung von Symptom-Checkern angegeben. Die vollständigen Angaben zu diesen qualitativen Studien finden sich im Appendix in Tabelle A-4.

Erfahrungen:
bisher wenig Verwendung;
Symptomeingabe
strukturiert,

hohe Bedienbarkeit,

verbale Interaktion fehlt,

Symptom-Beschreibung
lückenhaft

3.4 Pilotevaluierung des Prozesses anhand einer ausgewählten DiGA („Symptoma“)

Im Anschluss an die soeben durchgeführte systematische Übersichtsarbeit zu Symptom-Checkern folgt nun ein Detail-Assessment zu einer einzelnen DiGA aus der Gruppe der Symptom-Checker:

**Detail-Assessment
einer DiGA**

Anhand des im Projektbericht 2020 [6] skizzierten Prozesses zur Bewertung von DiGA, welcher im Hintergrundkapitel 1.2 des vorliegenden Berichts beschrieben ist, soll in diesem Abschnitt erhoben werden, wie die vorliegende Evidenz einer ausgewählten DiGA („Symptoma“) aus der Gruppe der Symptom-Checker im Abgleich mit den geforderten Evidenz- und Evaluationsanforderungen einzuordnen ist. Die Auswahl des Symptom-Checkers „Symptoma“ erfolgte vor Projektbearbeitung des AIHTAs durch den Auftraggeber (vgl. Methodenkapitel 2.4). Zunächst erfolgt ergänzend zum skizzierten Prozess, entsprechend den Ergebnissen der ersten beiden Forschungsfragen des aktuellen Berichts, die Kategorisierung und Bestimmung der ASVG-Relevanz sowie eine Überprüfung der technologischen Anforderungen an Interoperabilität und ELGA-Kompatibilität.

**Pilotevaluierung
des Prozesses anhand
DiGA „Symptoma“
(Symptom-Checker)**

3.4.1 Kategorisierung von „Symptoma“ und Bestimmung der ASVG Relevanz

Nach Anwendung der in Kapitel 3.1.1 beschriebenen Kategorisierungs- und Priorisierungskriterien ist die DiGA „Symptoma“ folgendermaßen zu klassifizieren:

- *Anwendungsbereich:* keine Einschränkung auf bestimmte Diagnosen (diagnose-unspezifischer Symptom-Checker).
- *Zielgruppe („Vulnerabilität“ der Patient*innen):* keine Zuordnung auf einen bestimmten Gesundheitszustand möglich (vgl. Kapitel 3.1.1: eine DiGA, welche ihre Zielgruppe nicht klar bestimmt, z. B. eine DiGA, die „alle Menschen“ mit oder ohne medizinisch relevante Symptome unterstützen soll, ist im Zweifelsfall der Zielgruppe „0-Gesunde ohne Risikofaktoren“ zuzuordnen).
- *Funktion:* Erkennung/Diagnose, Diagnosestellung (Erstellung von Diagnosevorschlägen durch den Symptom-Checker).

**Anwendungsbereich/
Zielgruppe:
keine Einschränkung**

**Funktion:
Erkennung/Diagnose**

**Anwender*in:
Nutzer*in und/oder
Leistungserbringer**

**keine unmittelbare
ASVG-Relevanz**

- *Anwender*innen*: Nutzer*innen (potentielle Patient*innen) und/oder Gesundheitspersonal
- *ASVG-Relevanz*: Aufgrund der Zielgruppe und der Funktion der DiGA „Symptoma“ keine unmittelbare ASVG-Relevanz (vgl. Kapitel 3.1.3).

3.4.2 Überprüfung der technologischen Anforderungen

technologie-spezifische Anforderungen

Zu den in Kapitel 3.2.1 beschriebenen technologie-spezifischen Aspekten wurden für die DiGA „Symptoma“ von der Website [8] folgende Angaben erhoben (vgl. Tabelle 3-11):

- CE-Kennzeichnung: vorhanden
- Medizinprodukt-Risikoklasse laut Hersteller: keine Angabe⁸
- Datenschutz: Datenschutzerklärung gemäß Datenschutz-Grundverordnung vorhanden und Datenspeicherung innerhalb EU
- Interoperabilität: keine Angabe
- ELGA-Kompatibilität: keine Angabe
- Algorithmen: keine Detailangabe⁹

3.4.3 Überprüfung der Risikoklasse nach MDR und Abgleich mit Evidenzanforderungen

CE-Kennzeichnung und Risikoklasse

CE-Kennzeichnung vorhanden; keine Einteilung der Risikoklasse nach EU MDR 2017/745

Das Medizinprodukt „Symptoma Digital Health Assistant“ verfügt über eine CE-Kennzeichnung [8]. Aktuell befindet sich die DiGA in Version 2,1 (Stand: 17. September 2021). Laut Hersteller fällt die DiGA in Medizinprodukt Risikoklasse 1 gemäß Richtlinie 93/42/EWG. Seit Mai 2021 wurde die Medizinprodukt Richtlinie 93/42/EWG durch die neue Medizinprodukt Richtlinie EU MDR 2017/745 abgelöst, nach welcher für die meisten DiGAs eine Einteilung in eine höhere Risikoklasse erfolgt. Für die neue Richtlinie EU MDR 2017/745 liegt für „Symptoma“ noch keine Einteilung der Risikoklasse des Herstellers vor.

mind. Risikoklasse 2a; Gefahr durch Verunsicherung/Folgeuntersuchungen/Fehleinschätzung der Dringlichkeit

Die DiGA „Symptoma“ ist laut EU MDR 2017/745 zumindest Risikoklasse 2a zuzuordnen (vgl. Methodenkapitel 2.4, Tabelle 2-2). Bei Aufrufen des „Symptoma“ Symptom-Checkers findet sich der Hinweis, dass die DiGA nicht unter 18 Jahren, bei Schwangerschaft, im Falle eines medizinischen Notfalls oder als Ersatz für einen ärztlichen Rat oder Diagnose verwendet werden soll. Für Patient*innen ist es jedoch nicht in jedem Fall möglich abzuschätzen, ob es sich um einen medizinischen Notfall handelt. In den generierten Diagnosevorschlägen finden sich auch Diagnosen, die eindeutig medizinische Notfälle darstellen. Mögliche Gefahren nach Verwendung des Symptom-Checkers sind Verunsicherung der Patient*innen mit potentiell unnötigen Folgeuntersuchungen und Eingriffen. Auch ein potentiell inadäquates Wiegen in Sicherheit der Patient*innen mit Ausbleiben einer Kontaktaufnahme zu geschultem Gesundheitspersonal stellt ein Risiko dar.

⁸ Laut Hersteller Medizinprodukt-Risikoklasse nach 93/42/EEC: Risikoklasse 1. Diese Medizinproduktregulierung wurde im Mai 2021 von der EU MDR 2017/745 abgelöst.

⁹ Verwendung der sog. „Symptoma Core“ als Kernelement, gestützt durch Algorithmen, medizinische Daten und künstliche Intelligenz.

Abgleich des vorliegenden Studiendesigns mit Evidenzstufen

Nach Anwendung des NICE Evidence Standards Framework (Neuversion: April 2021 veröffentlicht [11]), ist der Symptom Checker „Symptoma“ in Evidenzstufe C einzuordnen, da sie den Diagnoseprozess unterstützt (vgl. Tabelle 3-10):

**NICE:
Evidenzstufe „C“**

Tabelle 3-10: Evidenzstufen für digitale Gesundheitsanwendungen

Evidenzstufen nach NICE Evidence Standards Framework [11]
<p>Stufe A Digitale Gesundheitsanwendungen ohne individuell messbare Effekte auf die Gesundheit von Anwender*innen, organisatorische Anwendungen auf Systemebene. <i>Beispiele:</i> elektronische Gesundheitsakte, Krankenhaus-Informationssysteme.</p>
<p>Stufe B Digitale Gesundheitsanwendungen, welche Gesundheitsinformationen anbieten, einfache Monitoringfunktionen übernehmen oder Anwendungen zur Kommunikation. <i>Beispiele:</i> Anwendungen mit Empfehlungen für einen gesunden Lebensstil, digitaler Kopfschmerzkalender, Anwendungen für Video-Sprechstunden mit Therapeut*innen.</p>
<p>Stufe C Digitale Gesundheitsanwendungen die der Prävention und dem Selbstmanagement von Krankheiten dienen. Digitale Gesundheitsanwendungen mit individuell messbaren Effekten auf die Gesundheit von Anwender*innen. Sie können parallel zu konventioneller Therapie eingesetzt werden. <i>Beispiele:</i> Anwendungen zur Raucherentwöhnung oder Gewichtsreduktion. Digitale Gesundheitsanwendungen, die der Diagnostik und Therapie von Krankheiten dienen. Digitale Gesundheitsanwendungen mit klinischen Konsequenzen (durch aktives Monitoring oder Berechnungen). Digitale Gesundheitsanwendungen mit individuell messbaren Effekten auf die Gesundheit von Anwender*innen. <i>Beispiele:</i> Anwendungen, die anhand eingegebener Symptome Diagnosevorschläge macht, Anwendungen die kognitive Verhaltenstherapie für Angststörungen anbieten, Anwendungen die mit einem externen Sensor oder Implantat verknüpft sind und dessen Daten automatisch und kontinuierlich für Fernmonitoring gesendet werden.</p>

Abkürzungen: NICE: National Institute of Health and Care Excellence

Für die DiGA „Symptoma“ liegt als vorliegendes Studiendesign eine retrospektive Fallvignetten-Studie zur Messung der diagnostischen Genauigkeit vor. Als Komparator wurde mit den Einschätzungen anderer Symptom Checker verglichen. Die Kontrollgruppe entspricht somit nicht dem Standard-of-care (Einschätzung durch Gesundheitspersonal). Der geforderte „minimum evidence standard“ für die Evidenzstufe 3b nach NICE Evidence Standards Framework [11] sieht eine Verbesserung relevanter Endpunkte, gemessen mit einer qualitativ hohen Interventionsstudie vor. Der geforderte „best practice standard“ für diese Evidenzstufe ist das Vorliegen einer hoch qualitativen randomisiert-kontrollierten Studie – das vorliegende Studiendesign der DiGA „Symptoma“ entspricht demnach nicht dem durch das NICE-Bewertungsinstrument vorgegebenen „best practice standard“.

**retrospektive
Fallvignetten-Studie**

**vorhandener Komparator:
andere Symptom-Checker
(kein SoC)**

Anwendung Assessment Frameworks

Im skizzierten Prozess des Projektberichts 2020 [6] wurden zwei Bewertungsinstrumente für DiGAs ausgewählt: das „Module for Mobile Medical Applications“ [12] und das „Digi-HTA Framework“ [13]. Basierend auf diesen Bewertungsinstrumenten folgt eine Übersicht und Einteilung der technologspezifischen Charakteristika der digitalen Gesundheitsanwendung „Symptoma“, um die im vorausgegangenen Projekt ausgewählten Bewertungsinstrumente auf Praxistauglichkeit an einem Beispiel zu erproben.

**Projektbericht Teil 1:
Anwendung der
empfohlenen Bewertungs-
instrumente**

Tabelle 3-11: Technologiespezifische Charakteristika der DiGA „Symptoma“, basierend auf „Module for MMA“ und „Digi-HTA Framework“ [12, 13]

Digitaler Gesundheitsassistent & Symptom-Checker „Symptoma“ [16]	
Name der Intervention	Symptoma
Version	Version 2,1
Webbasiert und/oder Smartphone-App	<ul style="list-style-type: none"> ■ Webbasiert: www.symptoma.com ■ Android Play Store: „Anamnesis“ (early access App des Symptom-Checkers von Symptoma GmbH) ■ Apple App Store: keine App vorhanden
Art der Intervention	Symptom-Checker
CE-Kennzeichnung	✓
Datenschutzangaben	✓
Medizinprodukt-Risikoklasse nach 93/42/EEC	Risikoklasse 1
Medizinprodukt-Risikoklasse nach 2017/745 MDR	k.A.
Finanzierung	Kofinanziert durch das Programm „Horizon 2020“ der Europäischen Union unter der Fördernummer 830017
Zielsetzung der Intervention	Den Nutzer*innen mögliche Ursachen (Diagnosevorschläge) für angegebene Symptome anzuzeigen.
Allgemeine Funktionsweise	Der Symptom-Checker fordert die Nutzer*innen auf, Symptome in einen Chatbot einzugeben. Basierend auf den eingegebenen Symptomen fragt der Symptom-Checker die Nutzer*innen weitere Fragen im Sinne einer Anamnese. Abschließend werden vom Symptom-Checker Diagnosevorschläge präsentiert, priorisiert nach berechneter Wahrscheinlichkeit
Input	Texteingabe (Symptome)
Output	Textausgabe (Diagnosevorschläge)
Beschreibung der Benutzeroberfläche (User interface)	Symptoma ist ein Symptom-Checker mit einer Chatbot-Schnittstelle. Der Symptom-Checker kann online aufgerufen werden. Der Symptom-Checker ist derzeit in 34 verschiedenen Sprachen verfügbar. Im Logo des Symptom-Checkers wird angegeben „Symptoma – better diagnosis“.
Beschreibung der Funktionsweise im Detail	Nach dem Zugriff auf die Website werden die Nutzer*innen aufgefordert, durch Komma getrennte Symptome einzugeben. Der Symptom-Checker stellt daraufhin eine Reihe von Multiple-Choice-Fragen. Die Multiple-Choice-Fragen beziehen sich auf das Alter, das Geschlecht sowie auf zusätzliche Symptome. Als Ergebnis erstellt der Symptom-Checker eine digitale Patient*innenakte und eine Liste möglicher Ursachen (Diagnosevorschläge). Die Patient*innendatei enthält das Alter und das Geschlecht sowie die angegebenen, verneinten und unsicheren Symptomen der Patient*innen. In der Liste möglicher Ursachen werden mögliche Diagnosen für die gemeldeten Symptome aufgeführt, geordnet nach der berechneten Wahrscheinlichkeit. Wenn die Nutzer*innen auf eine mögliche Diagnose klicken, werden weitere Informationen zu dieser Diagnose angezeigt, sowie ein Link zur Informationsquelle.
Informationsquelle	Als Informationsquelle für die Diagnosevorschläge dienen eine Vielzahl an Websites. Für ein fiktives, beispielhaftes Symptom-Assessment (zum Leitsymptom „Kopfschmerzen“) wurden folgende Websites als Quellen angezeigt: <ul style="list-style-type: none"> ■ eref.thieme.de ■ netdokter.at ■ tekiatsu.at ■ flexikon.doccheck.com ■ vitanet.de ■ homoeopathie-online.info ■ schmerzlinik.de ■ stada.de ■ bz-berlin.de ■ gesundheit.de ■ vitamin-b12-mangel.org
Zugang zur digitalen Gesundheitsanwendung	Die DiGA wird über eine Website aufgerufen. Im Google Play Store ist eine App „Anamnesis“ (early access App des Symptom-Checkers von Symptoma GmbH) vorhanden. Im Apple App Store ist keine App vorhanden. Die Website verfügt über eine mobile Ansicht für Smartphones. Keine Einschränkung nach Betriebssystem.
Verwendeter Algorithmus	Verwendung der sog. „Symptoma Core“ als Kernelement, gestützt durch Algorithmen, medizinische Daten und künstliche Intelligenz. Keine Detail-Angabe
Software-Updates für Re-Evaluierung	Software-Updates mit Erweiterung der Funktionen (z. B. Triage-Empfehlungen) Software-Updates, welche die Informationsquelle für Diagnosevorschläge adaptieren.

Digitaler Gesundheitsassistent & Symptom-Checker „Symptoma“ [16]	
Kosten	Die DiGA ist kostenfrei verfügbar
Registrierung	Es ist keine Registrierung erforderlich
Evidenz Symptom-Checker	2 RCTs, 24 Beobachtungsstudien
Evidenz Symptoma	1 Beobachtungsstudie ¹⁰

Abkürzungen: CE: europäische Konformität („conformité européenne“); k.A.: keine Angabe; MDR: Medizinprodukteverordnung („Medical Device Regulation“); MMA: digitale Gesundheitsanwendung („Mobile Medical Application“); RCT: randomisiert kontrollierte Studie („randomized controlled trial“)

3.5 Prozesskonzeption für die Bewertung von digitalen Gesundheitsanwendungen

Im Folgenden werden die adaptierten Prozessschritte beschrieben. Diese lassen sich folgendermaßen einteilen:

- Identifikation und Einreichung der DiGA, welche für eine Refundierung in Betracht gezogen wird,
- Kategorisierungs- und Filterprozess der eingereichten DiGA,
- Überprüfung der Risikoklasse (nach Medizinprodukteverordnung), Evidenzanforderungen und technologiespezifischen Anforderungen,
- Nutzenbewertung durch HTA-Evaluation,
- Entscheidung über Refundierung.

**Adaptierung
der Prozessschritte**

Nach einer Beschreibung der einzelnen, empfohlenen Prozessschritte folgt ein Fragenkatalog zu den noch offenen Fragen, welche verfahrenstechnisch und politisch zu klären sind. Diese Prozessschritte und offenen Fragen sollten idealerweise in einem Workshop mit den Sozialversicherungsträgern auf Praxistauglichkeit diskutiert werden.

**Workshop zu
Praxistauglichkeit**

Identifikation/Einreichung der DiGA

Die Einreichung der identifizierten DiGA, welche für eine Refundierung in Betracht kommt, sollte in einem Antragsverfahren konzipiert sein. Eine Alternative stellen ko-kreative Entwicklungen von Anwendungen (in Kooperation mit den Sozialversicherungsträgern) dar.

**Identifikation,
Antrag**

Nach Einreichungen: Filterprozess

Nach Einreichung einer DiGA sollte der Filterprozess gestartet werden. Dieser umfasst folgende Schritte:

**Filterprozess
(Priorisierungskriterien)**

1. Kategorisierung der DiGA nach Zielsetzung, Funktion und Anwendergruppe,
2. Überprüfung der ASVG Relevanz (Fokus auf therapeutische Funktionen für chronisch Erkrankte),
3. Überprüfung der technologischen Anforderungen (Angaben zum Datenschutz, zu den verwendeten Algorithmen, zu den technischen Standards für die Interoperabilität sowie der Kompatibilität mit ELGA).

¹⁰ Es wurden zusätzlich 2 Beobachtungsstudien für die COVID-19 diagnose-spezifische Variante von Symptoma identifiziert.

Überprüfung der Risikoklasse, Evidenzanforderungen und technologiespezifischen Anforderungen

Überprüfung (Risikoklasse und Evidenzstufe) Für DiGAs, welche die Filter- und Priorisierungskriterien erfolgreich durchlaufen, gilt es in einem nächsten Schritt die Risikoklasse (nach Medizinproduktverordnung 2017/745) und die Evidenzanforderungen (etwa nach dem NICE Evidence Standards Framework [11]) zu erheben. Anschließend wird ein Screening der vorliegenden Studien der DiGA durchgeführt.

Bei positivem Ergebnis des Filterprozesses: HTA-Evaluation

HTA-Evaluation, RoB-Bewertung, ggf. SR technologie-spez. Bewertung In einem nächsten Schritt folgt die Nutzenbewertung der DiGA durch eine HTA-Evaluation der Studien. Dabei kann gegebenenfalls eine systematische Übersichtsarbeit durchgeführt werden. Die Risk-of-Bias Bewertung der Studien ist jedenfalls notwendig. Die Nutzenbewertung ist mit technologie-spezifischen Bewertungen (Datenschutzangaben, verwendeten Algorithmen sowie verwendeten technischen Standards) zu ergänzen.

Entscheidung über Refundierung

Refundierungsentscheidung Für DiGAs, welche durch die HTA-Evaluation einen ausreichenden Nutzen nachweis erbracht haben, ist eine Entscheidung über das Refundierungsmodell (Preisgestaltung, Ausmaß und Dauer der Refundierung) erforderlich.

Offene Fragen

Fragenkatalog Während den Expert*innenkonsultationen und in der Feinkonzeption des Prozesses wurde deutlich, dass es für einige verfahrenstechnische Fragen Klärungsbedarf gibt. Diese verfahrenstechnischen Fragen sind in der nachfolgenden Tabelle 3-12, gruppiert nach betreffenden Prozessschritt, dargestellt:

Tabelle 3-12: Fragenkatalog zu den verfahrenstechnischen, offenen Fragen zum Bewertungsprozess für digitale Gesundheitsanwendungen

Fragenkatalog	
Identifikation und Einreichung	Wer kann eine DiGA für eine Refundierung/Evaluation einreichen (Hersteller oder Nutzergruppen)?
	Welche Unterlagen sind dafür erforderlich?
	Erfolgt die Einreichung bei den Sozialversicherungsträgern?
	Bedarf es rechtlicher Regelungen für DiGAs insbesondere im Sozialversicherungsrecht?
	Wie sind DiGAs im leistungsrechtlichen System der Sozialversicherung einzuordnen?
	Ist ein Anerkennungsprozess für DiGAs (analog zum Erstattungskodex-Prozess) notwendig?
	Wie können Kriterien für diese Anerkennung (etwa auch im Sinne einer „Positivliste“) aussehen?
	Wie ist mit gleichwertigen (aber kostengünstigeren) DiGAs umzugehen (vgl. Meta-Verzeichnis Tabelle A-6)?
	Wie sind institutionelle Zuständigkeiten (hinsichtlich Bewertungsboard etc.) aufzuteilen?
	Bedarf es eines Registers/Verzeichnisses für geprüfte DiGAs (vgl. Kapitel 3.1.1)?
Alternativ: Ko-kreative Entwicklung von DiGAs in Kooperation mit den Sozialversicherungen	Wer identifiziert die Bedarfe der Sozialversicherungen?
	In welchem Technologie-Entwicklungsstadium ist eine solche Kooperation sinnvoll?
Filterprozess	Sollte bei Antragsstellung eine Versorgungslücke (Bedarf) ein Auswahlkriterium sein?
	ELGA-Kompatibilität ein Kriterium für Filterauswahl?
	Wie ist mit DiGAs umzugehen, welche die Daten außerhalb der EU speichern?
	Wie ist mit DiGAs umzugehen, für welche keine Angaben zu verwendeten Algorithmen vorliegen?

Fragenkatalog	
Evidenzanforderungen	Sollen die Evidenzanforderungen des NICE Evidence Standards Frameworks [11] in Österreich zur Anwendung kommen?
	Bedarf es einer Festlegung der erwarteten Endpunkte (klinische und system-relevante) sowie der verwendeten Komparatoren?
	Ist „keine Therapie“ als Komparator für DiGAs gerechtfertigt (z. B. wenn eine Versorgungslücke besteht)?
HTA-Evaluation	Wie ist mit äquivalenten DiGAs in der HTA-Evaluation umzugehen?
	Wann ist eine systematische Übersichtsarbeit erforderlich?
	Welche Institution überprüft die Datenschutzangaben?
	Welche Institution überprüft die technischen Standards sowie die Kompatibilität mit ELGA?
	Ab wann ist ein Re-Assessment der DiGA erforderlich (iterative Software-Updates)?
Refundierung	Wie erfolgt die Preisgestaltung bzw. Preisfestlegung für DiGAs, für welche ein Nutznachweis erbracht wurde?
	Wie ist mit DiGAs umzugehen, welche in vergleichbaren Gesundheitssystemen bereits refundiert werden?
	Wie ist die Verrechnung konkret zu gestalten (über Anspruchsberechtigte und/oder Leistungserbringer)?

Abkürzungen: DiGA – digitale Gesundheitsanwendung; ELGA – elektronische Gesundheitsakte; EU – europäische Union; HTA – Health Technology Assessment; NICE – National Institute for Health and Care Excellence

Design einer möglichen Refundierung von digitalen Gesundheitsanwendungen in Österreich

Eine grafische Veranschaulichung der vorgeschlagenen Prozesskonzeption für die Erstattung von DiGAs in Österreich ist in der nachfolgenden Abbildung 3-3 dargestellt.

Als zwingende Bedingung werden im Prozess-Design die Art der Anwendung, die Sprache, die Priorisierung nach ASVG-Kriterien, das Einhalten der Datenschutzvorgaben entsprechend GDPR, das Vorliegen einer CE-Zertifizierung und ein positiver Nutznachweis vorgeschlagen.

Die Kompatibilität mit ELGA, der Umgang mit inadäquaten Studiendesigns/ Komparatoren und die Kriterien für ein Re-Assessment werden in Abhängigkeit der noch offenen verfahrenstechnischen Fragen als mögliche Bedingungen angeführt.

Die Prozesskonzeption sieht die Priorisierungskriterien entsprechend ASVG-Relevanz nach Antragsstellung zur Refundierung einer DiGA vor; gegebenenfalls können die Priorisierungskriterien bereits vor Antragsstellung als Filterkriterium eingesetzt werden.

Wir empfehlen den DiGA-Prozess nach Klärung der noch offenen verfahrenstechnischen und politischen Fragen entsprechend weiter zu adaptieren.

Prozesskonzeption für DiGAs in Österreich

zwingende und mögliche Bedingungen

Adaptierung nach Klärung der noch offenen Fragen

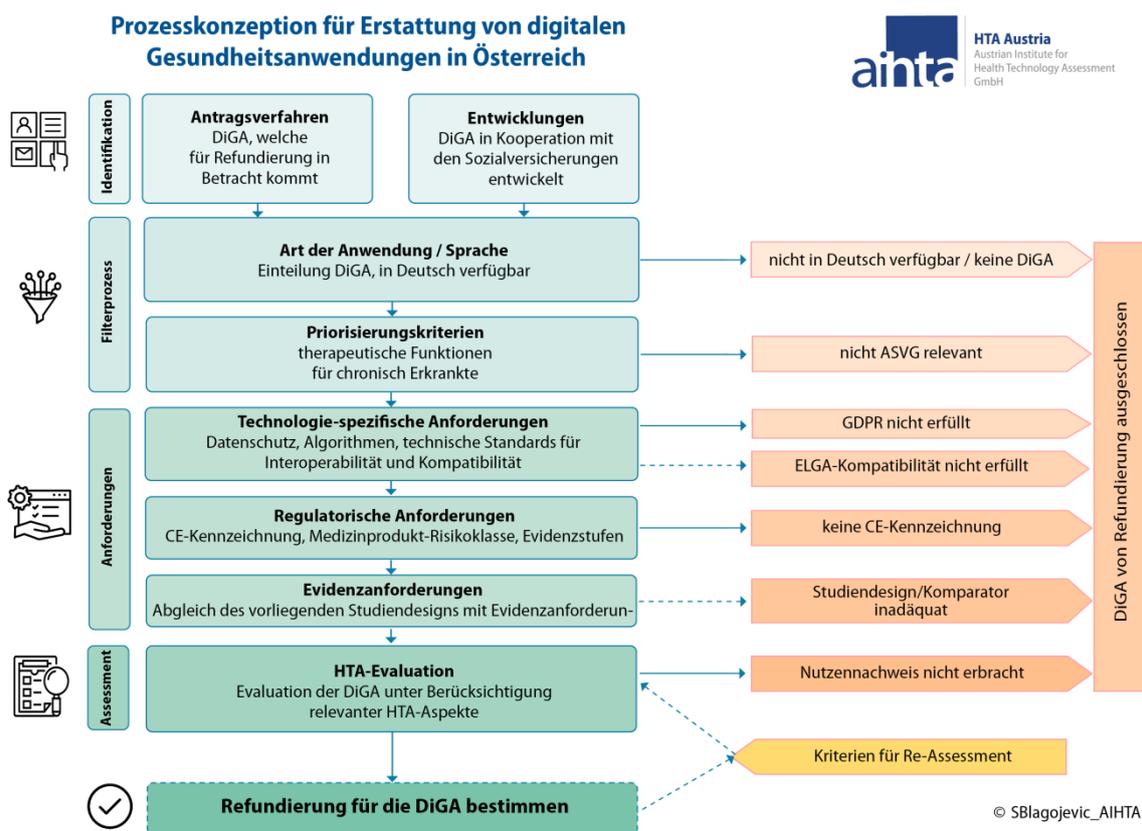


Abbildung 3-3: Prozesskonzeption für die Erstattung von digitalen Gesundheitsanwendungen in Österreich

4 Diskussion

Der vorliegende Bericht zielte darauf ab, eine verfeinerte praxistaugliche Prozesskonzeption für die evidenzbasierte Bewertung digitaler Gesundheitsanwendungen (DiGAs) vorzulegen. Dazu wurden DiGAs identifiziert, kategorisiert und priorisiert. In einem weiteren Schritt wurde eine systematische Übersichtsarbeit zu DiGAs aus der Gruppe der Symptom-Checker erstellt. Anschließend wurde eine Pilotevaluierung des im Vorjahr skizzierten Prozesses des AIHTA [6], anhand einer konkreten DiGA („Symptoma“) durchgeführt. Die daraus gewonnenen Erkenntnisse flossen in die Verfeinerung des für Österreich empfohlenen Bewertungsprozesses für DiGAs ein.

**Ziel der Arbeit:
praxistaugliche
Prozesskonzeption für die
evidenzbasierte
Bewertung von DiGAs**

4.1 Kategorisierung, Priorisierungskriterien und technologiespezifische Anforderungen

Zunächst wurde ein Meta-Verzeichnis bestehend aus 132 mobilen App- und Webanwendungen erstellt. Darauf aufbauend wurden diese – mit Hilfe eines bereits entwickelten Kategoriensystems – kategorisiert. Es ist des Weiteren zu konstatieren, dass es unter Anwendung der in diesem Bericht entwickelten – auf dem Allgemeinen Sozialversicherungsgesetz (ASVG) basierenden – Priorisierungskriterien (in Deutsch verfügbar, Einschränkung auf Chronikerprogramme mit Funktionen zur Behandlung und Monitoring) derzeit 38 DiGAs gibt, die für eine unmittelbare Evaluierung in Frage kämen. Die Indikationsbereiche dieser priorisierten DiGAs umfassen psychische Störungen (n=9), Endokrine Ernährungs- und Stoffwechselerkrankungen (n=7), Herz-Kreislaufkrankungen (n=7) und „andere“ Indikationen (n=15).

**Erstellung eines
Meta-Verzeichnisses:**

**Kategorisierung,
Priorisierung nach
ASVG-Relevanz**

**Unterstützung von
chronisch Kranken**

Zu den technologiespezifischen Charakteristika dieser priorisierten DiGAs konnte in diesem Bericht gezeigt werden, dass für den Großteil eine CE-Kennzeichnung (n=29) und Angaben zum Datenschutz (n=35) vorliegen. Eine detaillierte Prüfung der Datenschutzangaben (GDPR-Konformität, Speicherort der Daten) sollte jedenfalls durch entsprechende Expert*innen erfolgen. Transparente Angaben zu den Funktionsweisen der von den priorisierten DiGAs verwendeten Algorithmen wurden von keinem der DiGA-Hersteller angegeben. Die Kompatibilität der DiGAs mit der österreichischen elektronischen Gesundheitsakte (ELGA) ist derzeit noch nicht ausgereift. Derzeit ist eine Übermittlung mittels sog. „CDA-Dokumente“ als Übergangslösung möglich. Diese technische Schwierigkeit könnte gegebenenfalls durch die Übermittlung von Einzelwerten mit neuerem Standard (sog. „FHIR-Standard“) zukünftig überwunden werden. Erste Demos mit diesem neueren Standard sind für die nächsten 1-2 Jahre geplant. Eine technische Kompatibilität mit ELGA könnte vor einer Evaluierung ein weiteres Filterkriterium darstellen. Da dadurch jedoch eine Abhängigkeit vom jeweiligen Entwicklungsstand der ELGA entsteht, ist zu diskutieren ob eine Kompatibilität mit ELGA zwingend erforderlich ist.

**technologiespezifische
Charakteristika:**

**CE-Kennzeichnung,
Datenschutz,
Kompatibilität mit
bestehenden Standards**

Bei Einteilung der 38 priorisierten DiGAs in die Evidenzstufen des NICE Evidence Standards Bewertungsinstruments [11] wird deutlich, dass die ASVG-relevanten, priorisierten DiGAs allesamt in die höchste Evidenzstufe C (mit entsprechend hohen geforderten Evidenzstandards) einzuordnen sind.

**Abgleich von
Risikoeinschätzung
mit Evidenzstufen**

4.2 Pilotevaluierung des Bewertungsprozesses für DiGAs

Erprobung des Prozesses an der DiGA-Gruppe „Symptom-Checker“

In einem nächsten Schritt wurde der skizzierte Prozess des vorausgegangenen Projekts [6] für die Bewertung einer DiGA aus der Kategorie der Symptom-Checker angewendet. Bei der Kategorisierung des Symptom-Checkers wurde deutlich, dass die DiGA der Funktionsgruppe „Erkennung (Diagnose)“ zuzuordnen ist, und damit die ASVG-relevanten Priorisierungskriterien stringent betrachtet nicht erfüllt. Die Pilotevaluierung wurde dennoch durchgeführt, um den skizzierten Prozess probenhaft zu durchlaufen.

4.2.1 Evidenz der Symptom-Checker

NIHR-Übersichtsarbeit (2019, 27 Studien) wurde „updated“ (mit weiteren 14 Studien)

Für die Wirksamkeits- und Sicherheitsanalyse für DiGAs aus der Gruppe der Symptom-Checker wurde die Evidenz einer systematischen Übersichtsarbeit des britischen National Institute for Health Research (NIHR) [9], welche 2019 publiziert wurde und 27 Studien umfasste, mit der Evidenz von 14 Studien im Sinne eines Updates ergänzt: acht Studien für die klinische Wirksamkeit, die Sicherheit, die diagnostische Genauigkeit und die Genauigkeit der Triage-Empfehlungen sowie sechs Studien für die Patient*innensicht.

Sicherheit: keine möglichen Schäden identifiziert (6 Beobachtungsstudien);

Zur Sicherheit (Schäden durch falsch-positive/falsch-negative Diagnosevorschläge oder Triage-Empfehlungen) konnte keine Evidenz zu möglichen Schäden durch Verwendung der Symptom-Checker identifiziert werden (in sechs Beobachtungsstudien gemessen). Die klinische Wirksamkeit (Krankheitsdauer oder Schwere der Erkrankung) wurde in einem RCT berichtet, dabei wurde ein statistisch nicht signifikanter Unterschied zuungunsten der Symptom-Checker gefunden.

Wirksamkeit: geringfügig längere Krankheitsdauer, nicht stat.sign. (1 RCT)

Für die diagnostische Genauigkeit konnten elf Beobachtungsstudien identifiziert werden. Dabei wurde in einer Studie eine Sensitivität von 93,3 % und eine Spezifität von 12,9 % angegeben. Als indirekter Parameter wurde die Übereinstimmung mit klinischer Diagnose (in Top 3 Vorschlägen des Symptom-Checkers gelistet) gewählt, diese reichte von 14-84,3 %. Die höchste diagnostische Genauigkeit konnte in Studien mit fiktiven klinischen Fallvignetten erzielt werden. Durch dieses Studiensetting sind jedoch nur äußerst begrenzte Annahmen auf die diagnostische Genauigkeit in realen Situationen zulässig.

Genauigkeit der Diagnosevorschläge und Handlungs-Empfehlungen: keine zufriedenstellenden Ergebnisse

Die Triage-Einschätzungen der Symptom-Checker wurde in zwölf Beobachtungsstudien untersucht. Dabei wurde in einer Studie eine Sensitivität von 84 % und eine Spezifität von 74 % angegeben. In den anderen Studien wurden indirekte Parameter zur Übereinstimmung mit den Triage-Empfehlungen angegeben, diese reichten von 33-100 %. Dieser Endpunkt wurde zum Großteil anhand fiktiver klinischer Fallvignetten und ohne adäquate Studienpopulationen analysiert.

systemisch/organisatorische Endpunkte: Evidenz deutet auf Einflussnahme hin

Systemisch/organisatorische Endpunkte (Einfluss auf die Anzahl der Arzt-/Ambulanzbesuche) wurden in einem RCT und sechs Beobachtungsstudien berichtet. Im RCT wurde berichtet, dass in der Interventionsgruppe häufiger Telefonkontakte mit Gesundheitspersonal, jedoch seltener Arzt*in-/Ambulanzbesuche stattgefunden haben. Es ist jedoch anzumerken, dass der evaluierte Symptom-Checker dieses RCTs den Nutzer*innen empfahl, sich direkt an die Gesundheitsnummer („NHS Direct“) zu wenden, anstatt eine*n Arzt*in-/Ambulanz aufzusuchen. Die Beobachtungsstudien lieferten inkon-

Zu-/Abnahme von Arzt-/Ambulanz-besuchen: inkonsistente Ergebnisse

sistente Ergebnisse. Die Evidenz deutet darauf hin, dass Symptom-Checker die Patient*innenflüsse beeinflussen können – aufgrund von inkonsistenten Ergebnissen ist jedoch keine Aussage über eine Zu- oder Abnahme von Arzt-/Ambulanzbesuchen nach Verwendung von Symptom-Checkern möglich.

Die patient*innenrelevanten Endpunkte deuten darauf hin, dass die meisten Nutzer*innen bisher keine Erfahrung mit der Verwendung von Symptom-Checkern haben. Die Nutzer*innen zeigten sich nach Verwendung der Symptom-Checker mit der Benutzeroberfläche tendenziell zufrieden. Als Wegbereiter für die Verwendung von Symptom-Checkern wurde unter anderem angegeben, dass das Symptom-Assessment personalisierter und strukturierter als eine herkömmliche Internet-Recherche empfunden wird. Als Hindernisse für die Verwendung der Symptom-Checker wurden hingegen die beschränkte Möglichkeit der vollständigen Beschreibung von Symptomen und das Fehlen der verbalen Interaktion mit dem Gesundheitspersonal angegeben. Laut den Autor*innen der systematischen Übersichtsarbeit des NIHR-Berichts [9] deutet die Evidenz der Patient*innenperspektive darauf hin, dass die Verwendung von Symptom-Checkern drastisch zunehmen wird, wenn diese in das Gesundheitssystem integriert werden sollten.

patient*innen-relevante Endpunkte:
hohe Anwenderfreundlichkeit, aber:
verbale Interaktion fehlend

4.2.2 Weitere Erwägungen zu Symptom-Checkern

Die Zweckbestimmung der Symptom-Checker liegt in einer Unterstützung des Diagnoseprozesses und einer Verbesserung der Steuerung von Patient*innenflüssen.

Ziel: bessere Diagnosestellung und Patient*innenflüsse

Durch die analysierten Studien wurde deutlich, dass Symptom-Checker in der Regel von Patient*innen verwendet werden. Für diagnostische Überlegungen oder Einschätzungen der Triage-Stufe durch Gesundheitspersonal spielen Symptom-Checker kaum eine Rolle, hierfür werden (wenn überhaupt) „clinical decision support systems“ verwendet [18, 88]. In den beschriebenen Studien des NIHR-Berichts wurden die Symptom-Checker ausschließlich von Patient*innen verwendet [9]. In den identifizierten Studien der Update-Suche gab es lediglich eine Studie, in welcher die Anwendung des Symptom-Checkers durch Klinikpersonal (oder Hersteller) zur Diagnoseunterstützung erfolgte [76].

Anwender:
in Studien immer Patient*innen

mit einer Ausnahme:
Klinikpersonal (oder Hersteller)

Durch die Sensitivität und Spezifität der Genauigkeit der Diagnosevorschläge und den Triage-Empfehlungen wird deutlich, dass die Überdiagnostik ein generelles Problem der Symptom-Checker darstellt. Die Problematik der Überdiagnostik in der Medizin wird durch die „Choosing Wisely“ Initiative beleuchtet [89]. An dieser Stelle kann ein Projekt der Medizinischen Universität Graz und der Donau-Universität Krems genannt werden, das sich unter anderem dem Problem der Überdiagnostik widmet: „gemeinsam gut entscheiden“ [90].

Gefahr der Überdiagnostik

Die Gefahr durch Fehlinformation der Diagnosevorschläge von Symptom-Checkern ist ebenfalls gegeben. Die Quelle der Daten ist häufig nicht leitlinienbasiert. So wurden in den Diagnosevorschlägen des für die Pilotevaluierung ausgewählten Symptom-Checkers für ein beispielhaftes Assessment zum Leitsymptom „Kopfschmerz“ unterschiedliche Websites angegeben: neben medizinischen Verlagen wurden auch äußerst fragwürdige Websites, wie z. B. eine Homöopathie-Website angegeben. Es ist deutlich darauf hinzuweisen, dass diese Quellen definitiv nicht den Anspruch evidenzbasierter Information für Diagnosevorschläge erfüllen.

Gefahr durch Fehlinformation der Diagnosevorschläge

**bislang ungeklärt:
Einflüsse auf
Gesundheitskompetenz,
Haftungsfragen etc.**

Weitere Überlegungen betreffen noch nicht gänzlich geklärte organisatorische Aspekte (digitale Gesundheitskompetenz, Konnektivität), ethische Aspekte (Datenschutz), legistische Aspekte (Haftungsfragen bei Schäden) sowie einen möglichen Einfluss der Verwendung von Symptom-Checkern auf die Ärzt*innen/Patient*innenbeziehung.

4.2.3 Beurteilung der Evidenz

**hohe Heterogenität
der Evidenz**

**Studientyp: Fallvignette,
Rollenspielszenarien**

**niedrige Evidenz für
Verallgemeinerung zu
diagnostischer Genauigkeit**

Die inkludierten Studien wiesen eine hohe Heterogenität auf. Die Symptom-Checker wurden nur teilweise an realen, ausgewählten Patient*innenfällen getestet. Sehr häufig wurde in neuen Studientypen auf das Studiensetting mittels Analyse von fiktiven klinischen Fallvignetten im Rollenspielszenario gesetzt. Ergebnisse zur Genauigkeit der Diagnosevorschläge und Triage-Empfehlungen bei klinischen Fallvignetten sind kritisch zu betrachten: einerseits ist die Übertragbarkeit auf reale Szenarien sehr eingeschränkt; andererseits kann nicht ausreichend ausgeschlossen werden, dass die Algorithmen der untersuchten Symptom-Checker mit denselben Fallvignetten trainiert wurden, welche für die Erhebung der diagnostischen Genauigkeit und Triage-Empfehlungen herangezogen wurden. Für die diagnostische Genauigkeit und die Triage-Empfehlungen konnten dadurch bisher keine zufriedenstellenden Ergebnisse geliefert werden.

**hohes
Verzerrungspotential**

Insgesamt wurde die Qualität der Evidenz der inkludierten Studien des NIHR-Berichts [9] als niedrig eingestuft (vorwiegend Beobachtungsstudien mit Verzerrungspotenzial). Die über die Update-Suche identifizierten Studien für die klinische Wirksamkeit, die Sicherheit, die Genauigkeit der Diagnosevorschläge und der Triage-Empfehlungen, wurden mit dem QUADAS-II Risk-of-Bias Tool analysiert. Dabei wurden sieben Studien mit hohem Verzerrungspotenzial und eine Studie mit niedrigem Verzerrungspotenzial eingestuft.

**unzureichender
Nutzenbeleg**

Im konkreten Fall der Symptom-Checker ist der Nutznachweis jedenfalls nicht ausreichend erbracht: weder für Patient*innen als Anwender, noch für die Unterstützung von Triage-Diensten für Gesundheitspersonal als Anwender.

4.3 Limitationen der Arbeit

**Anwendung von
Kategorisierungsansatz,
auch andere vorhanden**

Zum heutigen Zeitpunkt gibt es noch keinen universalen Kategorisierungsansatz für die Einteilung der DiGAs. Während der in diesem Bericht gewählte Kategorisierungsansatz für die Erstellung des DiGA Meta-Verzeichnisses gut geeignet war, ist darauf hinzuweisen, dass es noch weitere Kategorisierungsansätze gibt [91]. Da die Kategorisierungsansätze jedoch in der Kernidee einander ähnlich sind, ist es unwahrscheinlich, dass die Verwendung eines anderen Kategorisierungsansatzes die Ergebnisse in diesem Bericht bedeutend verändert hätte.

**NIHR-Bericht nur zu
Patient*innen als Anwender**

**Update auch zu Triage-
Diensten als Anwender**

Eine Limitation der systematischen Übersicht über die Symptom-Checker ist die Abweichung der in diesem Bericht gewählten PICO-Fragestellung zur systematischen Übersichtsarbeit des NIHR-Berichts [9]. Die Ein- und Ausschlusskriterien wurden im NIHR-Bericht nicht ausreichend klar beschrieben. Nach Durchsicht der eingeschlossenen Studien ist davon auszugehen, dass als Population/Anwender der Symptom-Checker nur Patient*innen ausgewählt wurden. Für diesen Bericht wurden in den PICO-Kriterien sowohl

Patient*innen als auch Gesundheitspersonal als Anwender definiert. Etwaige Studien über Symptom-Checker mit der Anwendergruppe Gesundheitspersonal, welche vor 2018 publiziert wurden, sind in dieser systematischen Übersichtsarbeit dadurch nicht berücksichtigt.

Eine weitere Limitation ist der Ausschluss diagnose-spezifischer Symptom-Checker (wie z. B. COVID-19 Symptom-Checker) für den vorliegenden Bericht; diese diagnose-spezifischen Symptom-Checker waren jedoch in der systematischen Übersichtsarbeit des NIHR-Berichts inkludiert [9]. Auch innerhalb der allgemeinen (diagnose-unspezifischen) Symptom-Checker, welche für die Synthese der Evidenz herangezogen wurden, gibt es Unterschiede in Bezug auf den Funktionsumfang. Auch hier erwarten wir keine Verzerrung in der Darstellung der Ergebnisse.

**Ausschluss
diagnose-spezifischer
Symptom-Checker**

5 Schlussfolgerungen

Im Folgenden fassen wir unsere Schlussfolgerungen in Bezug auf den Bewertungsprozess von digitalen Gesundheitsanwendungen zusammen:

- Für eine zielgerichtete, zeiteffiziente Bewertung von digitalen Gesundheitsanwendungen bedarf es einer Kombination aus Kategorisierungs- und Priorisierungskriterien, um aus der Fülle der verfügbaren digitalen Gesundheitsanwendungen diejenigen auszuwählen, die eine unmittelbare ASVG-Relevanz bedingen. Diese Priorisierungskriterien könnten entweder vor, jedoch spätestens nach Antragsstellung von DiGAs angewendet werden. Nach Anwendung dieser Filterkriterien auf die digitalen Gesundheitsanwendungen ausgewählter Verzeichnisse und Listungen wäre, zum jetzigen Stand, für 38 DiGAs das Durchlaufen des weiteren Bewertungsprozesses möglich. Die Kompatibilität mit ELGA könnte vor einer Evaluierung ein weiteres Filterkriterium darstellen.
- DiGAs aus der Gruppe der Symptom-Checker stellen nach Anwendung der Filterkriterien keine digitalen Gesundheitsanwendungen mit unmittelbarer ASVG-Relevanz dar, da sich Symptom-Checker in der Regel nicht an chronisch erkrankte Patient*innen richten. Ein weiterer Ausschlussgrund stellt die Einteilung der Funktion dar (diagnostische statt therapeutische Funktion).
- Bisher konnte für Symptom-Checker kein ausreichender Nutznachweis erbracht werden. Ein Problem der Symptom-Checker stellt eine mögliche Überdiagnostik dar. Gleichzeitig bestehen Bedenken in Bezug auf Fehlinformation bei falschen Diagnose- und Triage-Empfehlungen. Für ein Re-Assessment sollten Nutzenbelege für Triage-Dienste als Anwender erwogen werden. Die Studien sollten jedenfalls in realen Situationen (ohne Verwendung von ausgewählten, fiktiven Fallvignetten) durchgeführt werden.
- Die Feinkonzeption des Prozesses zur Bewertung von digitalen Gesundheitsanwendungen wurde entsprechend der Erkenntnisse der Pilotevaluierung des Prozesses adaptiert. In der gründlichen Evaluation von digitalen Gesundheitsanwendungen (insbesondere der strengen Datenschutzvorgaben und der Kompatibilität mit den technischen Schnittstellen zur elektronischen Gesundheitsakte) ist die Zusammenarbeit verschiedener österreichischer Institutionen (etwa ELGA GmbH) mit entsprechender Aufgabenverteilung erforderlich.
- In Bezug auf die Prozessschritte der digitalen Gesundheitsanwendungen gibt es offene Fragestellungen, die verfahrenstechnisch und politisch zu klären sind. Die Prozessschritte sollten idealerweise in einem Workshop mit den Sozialversicherungsträgern auf Praxistauglichkeit diskutiert werden.

Kategorisierungs- und Priorisierungskriterien zur Selektion:

Prüfung der ASVG-Relevanz

Symptom-Checker: keine ASVG-Relevanz

große Gefahr der Überdiagnostik und von Fehlinformation

weitere Prozesskonzeption in Zusammenarbeit mit DiGA-relevanten österreichischen Institutionen

offene Fragen: von/mit Sozialversicherung zu lösen

6 Literatur

- [1] Kramer U., Borges U., Fischer F., Hoffmann W., Pobiruchin M. and Vollmar H. C. [DNVF-Memorandum – Health and Medical Apps]. *Gesundheitswesen*. 2019;81(10):e154-e170. Epub 2019/10/02. DNVF-Memorandum – Gesundheits- und Medizin-Apps (GuMAs). DOI: 10.1055/s-0038-1667451.
- [2] Bundesministerium für Gesundheit. Digitale Gesundheitsanwendungen. 2021 [cited 11.10.2021]. Available from: <https://www.bundesgesundheitsministerium.de/themen/krankenversicherung/online-ratgeber-krankenversicherung/arzt-heil-und-hilfsmittel/digitale-gesundheitsanwendungen.html>.
- [3] Jeindl R. and Wild C. Technologiebewertung digitaler Gesundheitsanwendungen für Refundierungsentscheidungen. *Wiener Medizinische Wochenschrift*. 2021. DOI: 10.1007/s10354-021-00881-3.
- [4] Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM). Das Fast Track Verfahren für digitale Gesundheitsanwendungen (DiGA) nach § 139e SGB V. 2020 [cited 01.10.2021]. Available from: https://www.bfarm.de/SharedDocs/Downloads/DE/Service/Beratungsverfahren/DiGA-Leitfaden.pdf?__blob=publicationFile.
- [5] Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM). DiGA-Verzeichnis. 2021 [cited 30.09.2021]. Available from: <https://diga.bfarm.de/de>.
- [6] Jeindl R. and Wild C. Framework zur Unterstützung von Refundierungsentscheidungen zu digitalen Gesundheitsanwendungen (mHealth) und dessen (retrospektive) Anwendung an ausgewählten Beispielen. AIHTA Projektbericht Nr.: 134. Wien: HTA Austria – Austrian Institute for Health Technology Assessment GmbH.: 2020 [cited 20.09.2021]. Available from: https://eprints.aihta.at/1279/1/HTA-Projektbericht_Nr.134.pdf.
- [7] Tarricone R., Petracca F., Ciani O. and Cucciniello M. Distinguishing features in the assessment of mHealth apps. *Expert Rev Pharmacoecon Outcomes Res*. 2021;21(4):521-526. Epub 2021/03/06. DOI: 10.1080/14737167.2021.1891883.
- [8] Symptoma GmbH. Medizinprodukt Symptoma. 2021 [cited 20.09.2021]. Available from: <https://www.symptoma.de/de/medical-device>.
- [9] Chambers D., Cantrell A., Johnson M., Preston L., Baxter S. K., Booth A., et al. Digital and online symptom checkers and assessment services for urgent care to inform a new digital platform: a systematic review. *Health Services and Delivery Research*, 2019 August 2019 [cited 26.08.2021]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK545124/>.
- [10] Medical Device Coordination Group (MDCG). Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 – MDR and Regulation (EU) 2017/746 – IVDR. 2019 [cited 30.09.2021]. Available from: <https://ec.europa.eu/docsroom/documents/37581>.
- [11] NICE National Institute for Health and Care Excellence. Evidence standards framework for digital health technologies. 2021 [cited 20.09.2021]. Available from: <https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies>.
- [12] Moshi M. R., Tooher R. and Merlin T. Development of a health technology assessment module for evaluating mobile medical applications. *Int J Technol Assess Health Care*. 2020;36(3):252-261. Epub 2020/05/19. DOI: 10.1017/s0266462320000288.
- [13] Haverinen J., Keränen N., Falkenbach P., Maijala A., Kolehmainen T. and Reponen J. Digi-HTA: Health technology assessment framework for digital healthcare services. *Finnish Journal of eHealth and eWelfare*. 2019;11:326. DOI: 10.23996/fjhw.82538.
- [14] Semigran H. L., Linder J. A., Gidengil C. and Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *Bmj*. 2015;351:h3480. Epub 2015/07/15. DOI: 10.1136/bmj.h3480.
- [15] Coney A., Tolond S., Glowinski A., Marks B., Swift S. and Palser T. Accuracy of online symptom checkers and the potential impact on service utilisation. *PLOS ONE*. 2021;16(7):e0254088. DOI: 10.1371/journal.pone.0254088.
- [16] Symptoma GmbH. About Symptoma. 2021 [cited 27.09.2021]. Available from: <https://www.symptoma.com/en/about>.

- [17] Ada Health GmbH. About us. 2021 [cited 21.09.2021]. Available from: <https://ada.com/about/>.
- [18] Sutton R. T., Pincock D., Baumgart D. C., Sadowski D. C., Fedorak R. N. and Kroeker K. I. An overview of clinical decision support systems: benefits, risks, and strategies for success. *Npj Digital Medicine*. 2020;3(1):17. DOI: 10.1038/s41746-020-0221-y.
- [19] Sönnichsen A. Fluch oder Segen? Symptom-Checker und Diagnostik-Apps. 2019 [cited 30.09.2021]. Available from: https://www.ebm-netzwerk.de/de/medien/pdf/ebm-kolumne-09_19.pdf.
- [20] Richens J. G., Lee C. M. and Johri S. Improving the accuracy of medical diagnosis with causal machine learning. *Nature Communications*. 2020;11(1):3923. DOI: 10.1038/s41467-020-17419-7.
- [21] Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz. 1450 – die Gesundheitsnummer. 2021.
- [22] AIT – Austrian Institute of Technology. AIT – Austrian Institute of Technology, 2021 [cited 07.10.2021]. Available from: <https://www.ait.ac.at/>.
- [23] Dachverband der Sozialversicherungsträger. Österreichische Sozialversicherung. 2021 [cited 07.10.2021]. Available from: <https://www.sozialversicherung.at/cdscontent/?contentid=10007.846560&portal=svportal>.
- [24] INiTS Universitäres Gründerservice Wien GmbH. INITS – Vienna’s High-Tech Incubator. 2021.
- [25] ITSV GmbH. ITSV – Home. 2021 [cited 07.10.2021]. Available from: <https://www.itsv.at/cdscontent/?contentid=10007.697193&portal=itsvportal>.
- [26] ELGA GmbH. ELGA: Startseite. 2021 [cited 07.10.2021]. Available from: <https://www.elga.gv.at/>.
- [27] Shea B. J., Reeves B. C., Wells G., Thuku M., Hamel C., Moran J., et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008. DOI: 10.1136/bmj.j4008.
- [28] Whiting P. F., Rutjes A. W., Westwood M. E., Mallett S., Deeks J. J., Reitsma J. B., et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536. Epub 2011/10/19. DOI: 10.7326/0003-4819-155-8-201110180-00009.
- [29] Bertelsmann Stiftung. Digital-Health-Anwendungen für Bürger. 2016 [cited 11.10.2021]. Available from: <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/digital-health-anwendungen-fuer-buerger/>.
- [30] Busse R., Panteli D., Lantzsch H. and Stephani V. Wege zu einer besseren Implementierung von digitalen Gesundheitsanwendungen in die Gesundheitsversorgung der GKV (unpublished). 2020 [cited 30.09.2021]. Available from: https://www.mig.tu-berlin.de/menue/research/aktuelle_projekte/idiga/.
- [31] Rechtsinformationssystem des Bundes. Bundesrecht konsolidiert: Gesamte Rechtsvorschrift für Allgemeines Sozialversicherungsgesetz, Fassung vom 07.10.2021. 2021 [cited 07.10.2021]. Available from: <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10008147>.
- [32] World Health Organization (WHO). Classification of digital health interventions v1.0. 2018 [cited 11.10.2021]. Available from: <https://www.who.int/reproductivehealth/publications/mhealth/classification-digital-health-interventions/en/>.
- [33] Breyer E. Aktualisierung von Leistungskatalogen. Eine vergleichende Analyse von Refundierungsprozessen für ärztliche Leistungen in ausgewählten Ländern. HTA-Projektbericht 2008/22.: 2008 [cited 21.09.2021]. Available from: https://eprints.aihta.at/817/1/HTA-Projektbericht_Nr22.pdf.
- [34] FHIR Foundation. HL7 FHIR, Release 4. 2019 [cited 11.11.2021]. Available from: <http://hl7.org/fhir/>.
- [35] ELGA GmbH. ELGA-Gesamtarchitektur. 2017 [cited 01.10.2021]. Available from: https://www.elga.gv.at/fileadmin/user_upload/Dokumente_PDF_MP4/Technisches/ELGA_Gesamtarchitektur_2.30a.pdf.
- [36] ELGA GmbH. ELGA-Schulungsunterlagen – Textbausteine zur Ergänzung und Unterstützung eigener Schulungsunterlagen. 2016 [cited 01.10.2021]. Available from: https://www.elga.gv.at/fileadmin/user_upload/Dokumente_PDF_MP4/Technisches/ELGA_Basis_fuer_Schulungsunterlagen_V2.0.pdf.
- [37] ELGA GmbH. ILF: Telemonitoring-Episodenbericht Guide. 2021 [cited 01.10.2021]. Available from: https://wiki.hl7.at/index.php?title=ILF:Telemonitoring-Episodenbericht_Guide.

- [38] Cardiac Sense. Cardiac Sense – Watching Over You. 2021 [cited 30.09.2021]. Available from: <https://www.cardiacsense.com/>.
- [39] AIT Austrian Institute of Technology GmbH. Telegesundheitsdienste Tirol, Steiermark und VAEB. 2018 [cited 30.09.2021]. Available from: https://kit.ait.ac.at/wp-content/uploads/2018/06/TGD-Tirol-Steiermark-VAEB_V2.0-2018-03-30_AIT.pdf.
- [40] FibriCheck. FibriCheck – Monitor your heart rhythm, prevent strokes. 2021 [cited 30.09.2021]. Available from: <https://www.fibrichk.com/>.
- [41] heartfish GmbH. heartfish. 2021 [cited 30.09.2021]. Available from: <https://www.heartfish.io/>.
- [42] Rehappy GmbH. Rehappy – Die Nachsorge-App bei Schlaganfall. 2021 [cited 30.09.2021]. Available from: <https://www.rehappy.de/>.
- [43] SAVING d.o.o. Savvy ECG – Monitor the activity of your heart. Anytime and anywhere.: 2021 [cited 30.09.2021]. Available from: <http://www.savvy.si/en/>.
- [44] Abbott SA/NV. FreeStyle Libre Flash Glucose Monitoring System. 2021 [cited 30.09.2021]. Available from: <https://myfreestyle.be/fr/>.
- [45] Medtronic Belgium. Guardian Connect App. 2021 [cited 30.09.2021]. Available from: <https://guardianconnect.medtronic-diabetes.co.uk/>.
- [46] mySugr GmbH. The mySugr app. 2021 [cited 30.09.2021]. Available from: <https://www.mysugr.com/en/diabetes-app/>.
- [47] Vitadio s. r. o. Overcome diabetes with a personal companion. 2021 [cited 30.09.2021]. Available from: <https://vitad.io/>.
- [48] aidhere GmbH. zanadio – Die neue digitale Adipositas-Therapie. 2021 [cited 30.09.2021]. Available from: <https://zanadio.de/>.
- [49] University of Liverpool. Catch It – Making Sense of Your Moods. 2021 [cited 30.09.2021]. Available from: <https://www.liverpool.ac.uk/csd/app-directory/catch-it/>.
- [50] GAIA AG. deprexis – overcome depression effectively. 2021 [cited 30.09.2021]. Available from: <https://info.deprexis.com/>.
- [51] Sympatient GmbH. Invirto – Die Therapie gegen Angst. 2021 [cited 30.09.2021]. Available from: <https://invirto.de/>.
- [52] mynoise GmbH. Kalmeda – die mobile Tinnitus-Therapie auf Rezept. 2021 [cited 30.09.2021]. Available from: <https://www.kalmeda.de/>.
- [53] Mindable Health GmbH. Mindable – App auf Rezept für Panikstörung und Agoraphobie. 2021 [cited 30.09.2021]. Available from: <https://www.mindable.health/>.
- [54] Selfapy GmbH. Selfapy – Online-Kurs bei Depression. 2021 [cited 30.09.2021]. Available from: <https://www.selfapy.com/kurse/depression>.
- [55] mementor DE GmbH. Somnio – Das digitale Schlaftraining. 2021 [cited 30.09.2021]. Available from: <https://somnia.io/>.
- [56] GAIA AG. velibra – Angst ist behandelbar. 2021 [cited 30.09.2021]. Available from: <https://de.velibra.com/>.
- [57] GAIA AG. vorvida – für einen bewussten Umgang mit Alkohol. 2021 [cited 30.09.2021]. Available from: <https://de.vorvida.com/>.
- [58] ResMed. AirView – Support your patients from diagnosis through ongoing care. 2021 [cited 30.09.2021]. Available from: <https://www.resmed.com/en-us/healthcare-professional/products-and-support/monitoring-and-data-management/airview/>.
- [59] CANKADO Service GmbH. CANKADO PRO-React Onco – Der digitale Therapiebegleiter für Brustkrebspatientinnen und -patienten. 2021 [cited 30.09.2021]. Available from: <https://diga.cankado.com/>.
- [60] Comarch S.A. Comarch HomeHealth – System for effective Remote Patient Monitoring. 2021 [cited 30.09.2021]. Available from: <https://www.comarch.com/healthcare/products/comarch-homehealth/>.

- [61] Comunicare Solutions. Comunicare – Care companion. 2021 [cited 30.09.2021]. Available from: <https://www.comunicare.be/home/>.
- [62] GAIA AG. elevida – Das Online-Programm für Menschen mit Fatigue bei MS. 2021 [cited 30.09.2021]. Available from: <https://elevida.de/>.
- [63] Florio GmbH. florio HAEMO – Enabling life beyond haemophilia.: 2021 [cited 30.09.2021]. Available from: <https://florio-haemo.com/>.
- [64] FOLLOWHEALTH S.L. (HumanITcare). HumanITcare – We’re Delivering Healthcare From Home. 2021 [cited 30.09.2021]. Available from: <https://humanitcare.com/>.
- [65] icometrix. icompanion – the digital companion for people with neurological conditions and their physicians. 2021 [cited 30.09.2021]. Available from: <https://icompanion.ms/>.
- [66] Fosanis GmbH. Mika – Krebs-App auf Rezept. 2021 [cited 30.09.2021]. Available from: <https://www.mitmika.de/>.
- [67] MindMate Inc. Alzheimer’s app – Mindmate. 2021 [cited 30.09.2021]. Available from: <https://www.mindmate-app.com/>.
- [68] Newsenselab GmbH. M-sense – Migräne-App auf Rezept. 2021 [cited 30.09.2021]. Available from: <https://www.m-sense.de/>.
- [69] Varian Medical Systems Belgium. Noona – Varian. 2021 [cited 30.09.2021]. Available from: <https://www.varian.com/products/software/care-management/noona>.
- [70] Medtronic Belgium. Maela – die Benchmark-Lösung für das Remote-Patientenmanagement. 2021 [cited 30.09.2021]. Available from: <https://www.maela.fr/de/>.
- [71] Vaica Medical Ltd. Vaica – Medication Adherence Customized Solutions. 2021 [cited 30.09.2021]. Available from: <https://www.vaica.com/>.
- [72] Vivira Health Lab GmbH. Vivira – Deine App auf Rezept gegen Schmerzen. 2021 [cited 30.09.2021]. Available from: <https://www.vivira.com/>.
- [73] Baker A., Perov Y., Middleton K., Baxter J., Mullarkey D., Sangar D., et al. A Comparison of Artificial Intelligence and Human Doctors for the Purpose of Triage and Diagnosis. *Frontiers in Artificial Intelligence*. 2020;3:100.
- [74] Nateqi J., Lin S., Krobath H., Gruarin S., Lutz T., Dvorak T., et al. Vom Symptom zur Diagnose – Tauglichkeit von Symptom-Checkern. *HNO*. 2019;67(5):334-342. DOI: 10.1007/s00106-019-0666-y.
- [75] Gilbert S., Mehl A., Baluch A., Cawley C., Challiner J., Fraser H., et al. How accurate are digital symptom assessment apps for suggesting conditions and urgency advice? A clinical vignettes comparison to GPs. *BMJ Open*. 2020;10(12):e040269. DOI: 10.1136/bmjopen-2020-040269.
- [76] Ronicke S., Hirsch M. C., Türk E., Larionov K., Tientcheu D. and Wagner A. D. Can a decision support system accelerate rare disease diagnosis? Evaluating the potential impact of Ada DX in a retrospective study. *Orphanet Journal of Rare Diseases*. 2019;14(1):69. DOI: 10.1186/s13023-019-1040-6.
- [77] Trivedi S., Littmann J., Stempien J., Kapur P., Bryce R. and Betz M. A Comparison Between Computer-Assisted Self-Triage by Patients and Triage Performed by Nurses in the Emergency Department. *Cureus*. 2021;13(3):e14002-e14002. DOI: 10.7759/cureus.14002.
- [78] Verzantvoort N. C. M., Teunis T., Verheij T. J. M. and van der Velden A. W. Self-triage for acute primary care via a smartphone application: Practical, safe and efficient? *PLOS ONE*. 2018;13(6):e0199284. DOI: 10.1371/journal.pone.0199284.
- [79] Hill M. G., Sim M. and Mills B. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. *Medical Journal of Australia*. 2020;212(11):514-519. DOI: <https://doi.org/10.5694/mja2.50600>.
- [80] Schmieding M., Mörgeli R., Schmieding M., Feufel M. and Balzer F. Benchmarking triage capability of symptom checkers against that of medical laypersons: Survey study (Preprint). *Journal of Medical Internet Research*. 2020;23. DOI: 10.2196/24475.

- [81] Aboueid S., Meyer S., Wallace J. R., Mahajan S. and Chaurasia A. Young Adults' Perspectives on the Use of Symptom Checkers for Self-Triage and Self-Diagnosis: Qualitative Study. *JMIR Public Health and Surveillance*. 2021;7(1):e22637-e22637. DOI: 10.2196/22637.
- [82] Eccles A., Hopper M., Turk A. and Atherton H. Patient use of an online triage platform: a mixed-methods retrospective exploration in UK primary care. *The British journal of general practice: the journal of the Royal College of General Practitioners*. 2019;69(682):e336-e344. Epub 2019/03/25. DOI: 10.3399/bjgp19X702197.
- [83] Meyer A. N. D., Giardina T. D., Spitzmueller C., Shahid U., Scott T. M. T. and Singh H. Patient Perspectives on the Usefulness of an Artificial Intelligence-Assisted Symptom Checker: Cross-Sectional Survey Study. *Journal of Medical Internet Research*. 2020;22(1):e14679-e14679. DOI: 10.2196/14679.
- [84] Miller S., Gilbert S., Virani V. and Wicks P. Patients' Utilization and Perception of an Artificial Intelligence-Based Symptom Assessment and Advice Technology in a British Primary Care Waiting Room: Exploratory Pilot Study. *JMIR human factors*. 2020;7(3):e19713-e19713. DOI: 10.2196/19713.
- [85] Nadarzynski T., Miles O., Cowie A. and Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. *DIGITAL HEALTH*. 2019;5:205520761987180. DOI: 10.1177/2055207619871808.
- [86] Winn A. N., Somai M., Fergestrom N. and Crotty B. H. Association of Use of Online Symptom Checkers With Patients' Plans for Seeking Care. *JAMA network open*. 2019;2(12):e1918561-e1918561. DOI: 10.1001/jamanetworkopen.2019.18561.
- [87] Converse L., Barrett K., Rich E. and Reschovsky J. Methods of Observing Variations in Physicians' Decisions: The Opportunities of Clinical Vignettes. *J Gen Intern Med*. 2015;30 Suppl 3(Suppl 3):S586-594. Epub 2015/06/25. DOI: 10.1007/s11606-015-3365-8.
- [88] Kwan J. L., Lo L., Ferguson J., Goldberg H., Diaz-Martinez J. P., Tomlinson G., et al. Computerised clinical decision support systems and absolute improvements in care: meta-analysis of controlled clinical trials. *BMJ*. 2020;370:m3216. DOI: 10.1136/bmj.m3216.
- [89] Choosing Wisely Initiative. Choosing Wisely – Promoting conversations between patients and clinicians. 2021 [cited 01.10.2021]. Available from: <https://www.choosingwisely.org/>.
- [90] Donau-Universität Krems – Department für evidenzbasierte Medizin und Evaluation. gemeinsam gut entscheiden – Choosing Wisely Austria. 2021 [cited 01.10.2021]. Available from: <https://gemeinsam-gut-entscheiden.at/>.
- [91] Institut für Gesundheitsförderung und Prävention (IfGP). Gesundheits-Apps – Grundlagenpapier unter besonderer Berücksichtigung des Aspekts Gesundheitskompetenz. 2018 [cited 30.09.2021]. Available from: <https://www.sozialversicherung.at/cdscontent/load?contentid=10008.712894&version=1532955053>.
- [92] Agoria vzw/asbl and beMedTech vzw/asbl. Belgian platform for medical mobile applications – mHealthBELGIUM. 2021 [cited 30.09.2021]. Available from: <https://mhealthbelgium.be/>.
- [93] NHS UK – United Kingdom National Health Service. NHS Apps Library. 2021 [cited 30.09.2021]. Available from: <https://www.nhs.uk/apps-library/>.
- [94] AIT – Austrian Institute of Technology. Center for Health & Bioresources (unpublished). 2021 [cited 30.09.2021]. Available from: <https://www.ait.ac.at/en/about-the-ait/center/center-for-health-bioresources>.
- [95] INiTS Universitäres Gründerservice Wien GmbH. Health Hub Vienna Alumni. 2021 [cited 30.09.2021]. Available from: <https://healthhubvienna.at/>.
- [96] Berchtesgadener Land Wirtschaftsservice GmbH. ScaleUp4Europe – Health Science. 2021 [cited 30.09.2021]. Available from: <https://scaleup4.eu/scaleups/>.

Anhang

Qualitätsbewertung des NIHR Berichts mit dem AMSTAR-II Assessment Tool

Tabelle A-1: Bewertung der Qualität der systematischen Übersichtsarbeit des National Institute of Health Research [9] unter Anwendung des AMSTAR-II Bewertungsinstruments [27]

Author, year	National Institute for Health Research, 2019 [9]
1. Did the research questions and inclusion criteria for the review include the components of Population, Intervention, Comparator and Outcome (PICO)?	Yes
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	Yes
3. Did the review authors explain their selection of the study designs for inclusion in the review?	Yes
4. Did the review authors use a comprehensive literature search strategy?	Yes
5. Did the review authors perform study selection in duplicate?	Yes
6. Did the review authors perform data extraction in duplicate?	Yes
7. Did the review authors provide a list of excluded studies and justify the exclusions?	Partial Yes
8. Did the review authors describe the included studies in adequate detail?	Yes
9. Did the review authors use a satisfactory technique for assessing the risk of bias in individual studies that were included in the review?	Yes
10. Did the review authors report on the sources of funding for the studies included in the review?	No
11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?	Not applicable
12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	Not applicable
13. Did the review authors account for risk of bias in individual studies when interpreting/discussing the results of the review?	Yes
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	Yes
15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	Not applicable
16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	Yes
Overall Confidence	High
Reasoning	No critical flaws suspected. Limitation: The authors did not report on the sources of funding for the studies included, but not considered a critical flaw. The authors did not provide a list of individual excluded studies, but not considered a critical flaw.

Extraktionstabellen

Systematische Übersichtsarbeit des NIHR-Berichts

Tabelle A-2: Datenextraktionstabelle zur systematischen Übersichtsarbeit des NIHR-Berichts [9]

Autor*innen, Jahr,	Chambers et al., 2019 [9]
Land	UK
Finanzierung	Öffentliche Finanzierung
Studiendesign	Systematic Review
Qualität (AMSTAR-II)	Hoch
Intervention	Symptom-Checker
Komparator	NHS 111, keine Restriktion
Setting	Allgemeinbevölkerung mit „dringendem“ Gesundheitsproblem (akut und chronisch)
Suchzeitraum	Bis 2018
Eingeschlossene Studien, Qualität der Evidenz	27, niedrig (2 RCTs, 25 Beobachtungsstudien)
Klinische Wirksamkeit (1 RCT¹¹)	
Krankheitsdauer, Schwere der Erkrankung	11,3 vs. 10,9 Tage (n.s.), NA
Sicherheit (6 Beobachtungsstudien)	
Schäden durch falsch-positive/falsch negative Diagnosevorschläge oder Triage-Empfehlungen	Keine der Studien berichtete von „Sicherheitsproblemen“
Genauigkeit der Diagnosevorschläge (7 Beobachtungsstudien)	
Sensitivität	1 Studie: 93,3 %
Spezifität	1 Studie: 12,9 %
Positiv prädiktiver Wert	NA
Negativ prädiktiver Wert	NA
Übereinstimmung mit klinischer Diagnose	Diagnose in Top 3 Vorschlägen des Symptom-Checker gelistet (4 Studien): 14-84,3 %
Diagnosezeitpunkt	NA
Genauigkeit der Triage-Empfehlungen (6 Beobachtungsstudien¹²)	
Sensitivität	NA
Spezifität	NA
Positiv prädiktiver Wert	NA
Negativ prädiktiver Wert	NA
Sicherheit/Angemessenheit der Triage-Empfehlungen	Triage-Empfehlungen als akkurat empfunden (6 Studien): 33-100 %

¹¹ In einer Beobachtungsstudie werden zwar klinische Endpunkte erwähnt, jedoch keine quantitativen Ergebnisse beschrieben.

¹² In einer Beobachtungsstudie werden zwar die Angemessenheit der Triage-Empfehlungen erwähnt, jedoch keine quantitativen Ergebnisse beschrieben.

Autor*innen, Jahr,	Chambers et al., 2019 [9]
Systemisch/organisatorische Endpunkte (1 RCT, 6 Beobachtungsstudien¹³)	
Anzahl Arzt-/Ambulanzbesuche	<p style="text-align: center;">1 RCT:</p> <p>in der Interventionsgruppe häufiger Telefonkontakt mit Gesundheitspersonal (2,5 % vs. 1,3 %; multivariate RR 2,53, 95 % CI: 1,10-5,82; p=0,029), jedoch seltener Arzt-/Ambulanzbesuche (16,3 % vs. 20,0 %, multivariate RR 0,71, 95 % CI: 0,53-0,95; p=0,019)</p> <p style="text-align: center;">6 Beobachtungsstudien:</p> <ol style="list-style-type: none"> 1. Symptom-Checker im Vergleich zu Telefon-Triage: häufiger Selbst-Management empfohlen (40 % vs. 14 %), seltener Allgemeinmediziner*in (29 % vs. 60 %), vergleichbar für Notfallmedizin (23 % vs. 21 %). 2. Von den Symptom-Checker Assessments wurden 72 % von Allgemeinmediziner*innen mit Erfordernis eines Arzt-/Ambulanzbesuchs eingestuft. 3. Nutzer*innen mit Symptom-Checker im Vergleich zu keiner elektronischen Diagnoseunterstützung: häufiger Bedürfnis nach Arzt-/Ambulanzbesuch (F(1) = 6,81; p>0,01) oder Notaufnahme (F(1) = 9,68; p > 0,01). 4. 18 % der Nutzer*innen planten einen Arzt-/Ambulanzbesuch aber führten diesen nach Verwendung des Symptom-Checkers nicht durch 5. Symptom-Checker im Vergleich zur Telefon-Triage: Symptom-Checker Empfehlungen: Allgemeinmediziner*in: 40 %; Rückruf durch Gesundheitspersonal: 20 %; Selbstbehandlung: 18 %; Notruf/Notaufnahme: 20 %; andere Services: 2 %; Telefon-Triage Empfehlungen: Allgemeinmediziner*in 60 %; Notruf/Notaufnahme: 21 %; Selbstbehandlung: 14 %; andere Services: 5 %. 6. durch Symptom-Checker im Vergleich zu Allgemeinmediziner*in häufiger dringlichere Triage-Empfehlungen.
Patient*innenrelevante Endpunkte (2 Beobachtungsstudien)	
Lebensqualität, Zufriedenheit der Patient*innen	NA
Bedienbarkeit der digitalen Gesundheitsanwendung	NA
Haltung der Patient*innen gegenüber Symptom-Checker	<p>1 qualitative Studie: Nutzer*innen sind gegenüber Symptom Checker gewogen</p> <p>1 Survey: 57 % der Patient*innen würden Triage Empfehlung befolgen</p>

Abkürzungen: CI: Konfidenzintervall; NA: nicht verfügbar (not available) NHS: National Health Service; NIHR: National Institute of Health Research; RCT: randomisiert kontrollierte Studie; RR: Risk Ratio; UK; United Kingdom;

¹³ In zwei Beobachtungsstudien werden zwar systemisch/organisatorische Endpunkte erwähnt, jedoch keine Ergebnisse beschrieben.

Studien über Symptom-Checker aus der Update-Suche

Tabelle A-3: Datenextraktionstabelle zu zusätzlich identifizierten Studien zu Symptom-Checkern (Teil 1)

Autor*innen, Jahr, Quelle	Baker et al., 2020 [73]	Gilbert et al., 2020 [75]	Hill et al., 2020 [79]	Nateqi et al., 2019 [74]
Land	UK	Deutschland, UK	Australien	Österreich
Finanzierung	8 Autor*innen: Mitarbeiter*innen von Babylon Health mit potenziellem Interessenkonflikt. 1 Autor: ohne kommerzielle oder finanzielle Beziehung als potenzieller Interessenkonflikt.	Gefördert durch Ada Health GmbH. Das Forschungsteam der Ada Health GmbH hat Forschungsgelder von der Foundation Botnar und der Bill & Melinda Gates Foundation erhalten.	Unterstützt durch ein Research Grant des Australian Government Research Training Program	2 Autor*innen: Gründer und Inhaber von Symptoma GmbH Alle Ko-Autor*innen: Mitarbeiter*innen von Symptoma GmbH
Studiendesign	Prospektive Fallvignetten-Studie	Prospektive Fallvignetten-Studie	Retrospektive Fallvignetten-Studie	Retrospektive Fallvignetten-Studie
Intervention	Babylon Triage and Diagnostic System	8 Symptom Checker Apps	36 Symptom-Checker: 17 Symptom Checker mit Diagnose-Information 9 Symptom Checker mit Triage-Empfehlungen 10 Symptom Checker mit Diagnose-Information und Triage-Empfehlungen	24 Symptom-Checker
Komparator	Ohne Komparator Referenzstandard: Allgemeinmediziner*innen	Ohne Komparator Referenzstandard: Allgemeinmediziner*innen	NA	Ohne Komparator Referenzstandard: andere Symptom Checker
Forschungsfrage, Setting	Klinische Vignetten mit Rollenspielszenario: Scheinkonsultationen zwischen Patient*in und Ärzt*in oder dem Chatbot Patient*innen gespielt von Allgemeinmediziner*innen (einige davon Mitarbeiter*innen von Babylon). Vier Runden an aufeinanderfolgenden Tagen, in jeder Runde gab es bis zu vier „Patient*innen“ und vier „Ärzt*innen“	Symptom Checker und Allgemeinmediziner*innen beurteilen klinische Vignetten. Die klinische Vignetten dienen als Gold-Standard.	Symptom-Checker wurden mit Fallvignetten getestet und auf ihre diagnostische Genauigkeit sowie die Angemessenheit der Triage-Ratschläge bewertet.	Aktualisierung einer Studie, in welcher 19 Symptom-Checker anhand 45 Patient*innenvignetten zur diagnostischen Präzision getestet wurden [14]. Ergänzung mit 5 Symptom-Checkern (Symptoma, Ada, FindZebra, Mediktor und Babylon).
Teilnehmer*innen	7 Allgemeinmediziner*innen 10 Patient*innen (gespielt durch Allgemeinmediziner*innen)	7 Allgemeinmediziner*innen	NA	NA
Merkmale der Teilnehmer*innen	Als Berater für Babylon tätig: 2/7 Ärzt*innen 5/10 Patient*innen Zuvor Mitarbeiter*in bei Babylon: 1/10 Patient*in An der Entwicklung von Babylon Produkten beteiligt: 1/10 Patient*in	NA	NA	NA

Autor*innen, Jahr, Quelle	Baker et al., 2020 [73]	Gilbert et al., 2020 [75]	Hill et al., 2020 [79]	Nateqi et al., 2019 [74]
Anzahl der Fallvignetten	Durchschnittliche Fallvignetten pro Ärzt*in: 56,6 Fallvignetten durch Intervention: 100	200 klinische Fallvignetten	48 klinische Fallvignetten (1170 Diagnose-Test Fallvignetten, 688 Triage-Test Fallvignetten)	45 klinische Fallvignetten (für diagnostische Treffergenauigkeit)
Klinische Wirksamkeit				
Krankheitsdauer, Schwere der Erkrankung	NA	NA	NA	NA
Sicherheit				
Schäden durch falsch- positive/falsch negative Diagnosevorschläge oder Triage-Empfehlungen	NA	NA	NA	NA
Genauigkeit der Diagnosevorschläge				
Sensitivität	Durchschnitt Ärzt*innen: 83,9 % (95 % CI: 75,6 %-92,3 %) Babylon AI: 80,0 %	Ada: 99,0 % Babylon: 51,5 % Buoy: 88,5 % K Health: 74,5 % Mediktor: 80,5 % Symptomate: 61,5 % Your.MD: 23,5 % WebMD: 93,0 %	NA	NR ¹⁴
Spezifität	NA	NA	NA	NR ¹⁴
Positiv prädiktiver Wert	Durchschnitt Ärzt*innen: 43,6 % (95 % CI: 36,3 %-50,9 %) Babylon AI: 44,4 %	NA	NA	NA
Negativ prädiktiver Wert	NA	NA	NA	NA
Übereinstimmung mit klinischer Diagnose	NA	<i>Diagnose in Top 3:</i> Durchschnitt Ärzt*innen: 82,1 %±5,2 % Ada: 70,5 % Babylon: 32,0 % Buoy: 43,0 % K Health: 36,0 % Mediktor: 36,0 %	27 diagnostische symptom checker: Korrekte Diagnose an 1. Stelle in 421/1170 Fallvignetten (36 %; 95 % CI: 31 %-42 %) Korrekte Diagnose innerhalb Top 3 in 606 Fallvignetten (52 %; 95 % CI: 47 %-59 %) Korrekte Diagnose innerhalb Top 10 in 681 Fallvignetten (58 %; 95 % CI: 53 %-65 %) Symptom checker mit künstlicher Intelligenz: Korrekte Diagnose an 1. Stelle in 46 % der Fallvignetten (95 % CI: 40 %-57 %)	<i>Diagnose in Top 1; 3; 10:</i> Symptoma: 82,2 %; 100 %; 100 % Babylon: 47 %; 64 %; 64 % ADA: 56 %; 78 %; 82 % FindZebra: 22 %; 33 %; 60 % Mediktor: 22 %; 31 %; 38 %

¹⁴ Ergebnisse als Abbildung dargestellt.

Autor*innen, Jahr, Quelle	Baker et al., 2020 [73]	Gilbert et al., 2020 [75]	Hill et al., 2020 [79]	Nateqi et al., 2019 [74]
Übereinstimmung mit klinischer Diagnose (Fortsetzung)		Symptomate: 27,5 % WebMD: 35,5 % Your.MD: 23,5 %	<i>Symptom checker ohne künstliche Intelligenz:</i> Korrekte Diagnose an 1. Stelle in 32 % der Fallvignetten (95 % CI: 26 %-38 %)	
Diagnosezeitpunkt	NA	NA	NA	NA
Genauigkeit der Triage-Empfehlungen				
Sensitivität	NA	NA	NA	NA
Spezifität	NA	NA	NA	NA
Positiv prädiktiver Wert	NA	NA	NA	NA
Negativ prädiktiver Wert	NA	NA	NA	NA
Sicherheit/Angemessenheit der Triage-Empfehlungen	Vergleich mit den Empfehlungen einer unabhängigen Person ¹⁵ : Durchschnitt Ärzt*innen: 93,1 % (95 % CI: 89,6 %-96,6 %) Babylon AI: 97,0 %	<i>Sicherheit der Triage-Empfehlung</i> ¹⁶ : Durchschnitt Ärzt*innen: 97,0 %±2,5 % Durchschnitt Apps: 90,1 %±7,4 % Ada: 97,0 % Babylon: 95,1 % Symptomate: 97,8 % Your.MD: 92,6 % Buoy: 80,0 % K Health: 81,3 % Mediktor: 87,3 %	<i>19 symptom checker with triage advice:</i> Korrekte Empfehlung in 338/688 Fallvignetten (49 %; 95 % CI: 44 %-54 %). Korrekte Empfehlung für Notfallversorgung: 63 %; 95 % CI: 52 %-71 % Korrekte Empfehlung für dringende Versorgung: 56 %; 95 % CI: 52 %-75 % Korrekte Empfehlung für nicht dringende Versorgung: 30 %; 95 % CI: 11 %-39 % Korrekte Empfehlung für Selbstversorgung: 40 %; 95 % CI: 26 %-49 %	NA
Systemisch/organisatorische Endpunkte				
Anzahl Arzt-/Ambulanzbesuche	NA	NA	NA	NA
Patient*innenrelevante Endpunkte				
Lebensqualität, Zufriedenheit der Patient*innen	NA	NA	NA	NA
Bedienbarkeit der digitalen Gesundheitsanwendung	NA	NA	NA	NA
Haltung der Patient*innen gegenüber Symptom-Checker	NA	NA	NA	NA

¹⁵ Die Triage-Sicherheit wurde danach beurteilt, ob die vorgeschlagene Triage-Aktion als sicher angesehen wurde, auch wenn sie übervorsichtig war.

¹⁶ Verwendete Triage-Stufen: 1 = Notruf setzen; 2 = in Notfallsambulanz begeben; 3 = Allgemeinmedizin innerhalb von vier Stunden; 4 = Allgemeinmedizin noch am selben Tag; 5 = Allgemeinmedizin ohne Dringlichkeit; 6 = Selbstversorgung. Triage-Sicherheit wurde definiert als maximal eine Stufe weniger konservativ als die Gold-Standard-Vignettenempfehlung.

Tabelle A-3: Datenextraktionstabelle zu zusätzlich identifizierten Studien zu Symptom-Checkern (Teil 2)

Autor*innen, Jahr, Quelle	Ronicke et al., 2019 [76]	Schmieding et al., 2021 [80]	Trivedi et al., 2021 [77]	Verzantvoort et al., 2018 [78]
Land	Deutschland	Deutschland, USA	Kanada	Niederlande
Finanzierung	Finanzierung als nicht zutreffend angegeben. Zwei Autor*innen erhalten ein Gehalt von der Ada Health GmbH. Ein Autor hält Anteile an der Ada Health GmbH.	Keine Drittmittel; Förderung durch die Deutsche Forschungsgemeinschaft und den Open Access Publikationsfonds der Charité – Universitätsmedizin Berlin	Keine Drittmittel	Finanziert von Achmea (niederländische Krankenkasse)
Studiendesign	Fall-Kontroll Studie	Prospektive Fallvignetten-Studie	Prospektive vergleichende Beobachtungsstudie	Prospektive Cross-Sectional Studie
Intervention	Ada DX (Anwender*innen: Gesundheitspersonal)	15 Symptom-Checker: HMS Family Health Guide; Healthy Children; Steps2Care; Symptify; Symptomate; Drugs.com; FreeMD; Doctor Diagnose; Family Doctor; Early Doc; Isabel; NHS; Symcat; Healthwise; iTriage	Algorithmischer Fragebogen mit Algorithmus-generiertem Self-Triage-Score	Smartphone App „Should I see a doctor?“ (in Dutch: „moet ik naar de dokter?“)
Komparator	Ohne Komparator Referenzstandard: Zeit bis zur bestätigten Diagnose in der elektronischen Gesundheitsakte (ohne Einsatz der Intervention)	Ohne Komparator Referenzstandard: Durchschnittslaie (n = 91)	Ohne Komparator Referenzstandard: Gesundheitspersonal (Triage-Pfleger*in)	Ohne Komparator Referenzstandard: Telephone Triage
Forschungsfrage, Setting	Ambulanz für seltene inflammatorische Systemerkrankungen; Fälle mit einer bestätigten Diagnose einer seltenen Erkrankung und einem dokumentierten Diagnosedatum wurden eingeschlossen. Subjektive Priorisierung von Fällen mit diagnostischen Odysseen. Die Zeit bis zur Diagnose wurde berechnet als der Zeitraum zwischen dem ersten dokumentierten Besuch mit Symptomen und dem dokumentierten Besuch der Diagnose.	Laien ohne professionellen medizinischen Hintergrund wurden rekrutiert und bewerteten klinische Fallvignetten. Die Genauigkeit der Triage-Empfehlungen der Symptom Checker wurde mit den Triage-Empfehlungen der Laien verglichen.	Patient*innen, die in einer Notaufnahme eines Krankenhauses vorstellig wurden (ohne Krankenwagen), wurde eine Reihe einfacher Fragen gestellt, die anhand eines Algorithmus einen Triage-Score generierten. Dieser Triage-Score wurde mit der kanadischen Triage- und Akuitätsskala verglichen, die von Gesundheitspersonal (Triage-Pfleger*in) vergeben wurde.	Smartphone App mit integriertem Fragebogen zur Klarheit der App, Zufriedenheit mit der App sowie Absicht, den Empfehlungen der App zu folgen. Eine Auswahl der Teilnehmer*innen wurde von Gesundheitspersonal (Triage-Pfleger*in) angerufen, um zu bewerten, ob die Empfehlungen der App mit der Empfehlung des Triage-Anrufs übereinstimmen.
Teilnehmer*innen	93 Fälle (elektronische Gesundheitsakten)	NA	492	4,456 (integrierter Fragebogen) 126 (Auswahl für Telefonbefragung)
Merkmale der Teilnehmer*innen	weibliche Fälle: 58 männliche Fälle: 35 Fälle mit multiplen Diagnosen: 12/93 Durchschnittliche Zeit zur Diagnose (Monate): 17 (4-65) Durchschnittliche Zahl der Besuche: 5,70 (3-7)	Alter (Jahre), Median (Bereich): 37 (20-73) Männlich: 60 % Triage Erfahrung innerhalb der letzten 6 Monate: Verwendung eines Symptom Checker: 22 % Triage Entscheidung: 25 %	Durchschnittsalter (Jahre): 47 (SD 22) Männlich: 49,8 % Empfundenes Problem: Lebensbedrohlicher Zustand, Extremitäten bedrohender Zustand: 29,7 % Herz- oder Lungenproblem: 15,2 %	Alter (Jahre): 0-18: 24,9 % 19-65: 72,1 % 66+: 3 % Männlich: 34 % Empfundenes Problem: Unterleibsschmerzen: 8 %

Autor*innen, Jahr, Quelle	Ronicke et al., 2019 [76]	Schmieding et al., 2021 [80]	Trivedi et al., 2021 [77]	Verzantvoort et al., 2018 [78]
Merkmale der Teilnehmer*innen (Fortsetzung)	Durchschnittliches Alter bei Diagnose (Jahre): 44,3 (33-56) Durchschnittliches Alter bei Symptombeginn (Jahre): 40 (27-50)	Weder Verwendung Symptom Checker noch Triage Entscheidung: 69 % Ohne medizinische Ausbildung: 88 % Erste-Hilfe-Ausbildung: 12 %	Unterleibsschmerzen, gastrointestinales Problem: 18,1 % Problem des Bewegungsapparates: 14,8 % Andere: 22,2 %	Hautbeschwerden: 7,8 % Husten: 6,7 % Unfall, Verletzung: 5,1 % Kopfschmerz: 4,8 % Beschwerden der Genitalorgane: 4,8 %
Anzahl der Fallvignetten	NA	45 klinische Fallvignetten	NA	NA
Klinische Wirksamkeit				
Krankheitsdauer, Schwere der Erkrankung	NA	NA	NA	NA
Sicherheit				
Schäden durch falsch-positive/falsch negative Diagnosevorschläge oder Triage-Empfehlungen	NA	NA	NA	NA
Genauigkeit der Diagnosevorschläge				
Sensitivität	NA	NA	NA	NA
Spezifität	NA	NA	NA	NA
Positiv prädiktiver Wert	NA	NA	NA	NA
Negativ prädiktiver Wert	NA	NA	NA	NA
Übereinstimmung mit klinischer Diagnose	NA	NA	NA	NA
Diagnosezeitpunkt	<p>Korrekte klinische Diagnose mit ADA DX früher als klinische Diagnose: Unter Top 5 Vorschlägen: 53,8 % Top Vorschlag: 37,6 %</p> <p>Medianer Vorteil des korrekten Diagnosevorschlags im Vergleich zum Zeitpunkt der klinischen Diagnose: 3 Monate (50 %) für Top 5 Vorschläge 1 Monat (21 %) für Top Vorschlag</p> <p>Korrekte klinische Diagnose mit ADA zum Zeitpunkt des ersten dokumentierten Patientenbesuchs: Unter Top 5 Vorschlägen: 33,3 % Top Vorschlag: 16,1 %</p>	NA	NA	NA

Autor*innen, Jahr, Quelle	Ronicke et al., 2019 [76]	Schmieding et al., 2021 [80]	Trivedi et al., 2021 [77]	Verzantvoort et al., 2018 [78]
Genauigkeit der Triage-Empfehlungen				
Sensitivität	NA	NA	NA	Sensitivität: 84 % (CI: 74 %-91 %)
Spezifität	NA	NA	NA	Spezifität: 74 % (CI: 58 %-86 %)
Positiv prädiktiver Wert	NA	NA	NA	Positiv prädiktiver Wert: 88 % (CI: 79 %-94 %)
Negativ prädiktiver Wert	NA	NA	NA	Negativ prädiktiver Wert: 67 % (CI: 51 %-81 %)
Sicherheit/Angemessenheit der Triage-Empfehlungen	NA	<p>Durchschnittslaien: Mittlere Triagegenauigkeit: 60,9 % (SD 6,8 %; 95 % CI: 59,5 %-62,3 %) Erkennung von Notfällen: 67,5 % (SD 16,4 %; 95 % CI: 64,1 %-70,8 %) Erkennung von Nicht-Notfällen: 68,4 % (SD 13,8 %; 95 % CI: 65,6 %-71,2 %) Erkennung von Selbstbehandlung: 46,7 % (SD: 15,9 %; 95 % CI: 43,4 %-49,8 %) Falsch-positive Notfälle: 4,10 % (p=0,002)</p> <p>Symptom-Checker: Mittlere Triagegenauigkeit: 58 % (SD 12,8 %) Erkennung von Notfällen: 80,6 % (SD 17,9 %) Erkennung von Nicht-Notfällen: 58,5 % (SD 29,1 %) Erkennung von Selbstbehandlung: 30,6 % (SD 25,7 %) Falsch-positive Notfälle: 24,7 % (p=0,002)</p>	<p>Übereinstimmung der Triage-Empfehlungen: 27 %. Überschätzung der Dringlichkeit (durch Patient*innen): 65 % Unterschätzung der Dringlichkeit (durch Patient*innen): 8 % Einschätzung für Aufnahme (durch Patient*innen): 41 % Tatsächliche Aufnahme durch Gesundheitspersonal (Triage-Pfleger*in): 17 %</p>	<p>Nutzer*innen der App: Empfehlung für Klinikbesuch: 58 % Empfehlung zur Selbstbehandlung: 34 % Empfehlung „wait-and-see“: 8 % App Empfehlung entspricht Empfehlung des Triage-Anrufs: 81 % Dringlichkeit Überschätzung (durch app): 8 % Dringlichkeit Unterschätzung (durch app): 11 %</p>
Systemisch/organisatorische Endpunkte				
Anzahl Arzt-/Ambulanzbesuche	NA	NA	NA	NA
Patient*innenrelevante Endpunkte				
Lebensqualität, Zufriedenheit der Patient*innen	NA	NA	NA	NA
Bedienbarkeit der digitalen Gesundheitsanwendung	NA	NA	NA	<p>App Bewertung 'neutral' bis 'sehr klar': 87 % App Bewertung 'neutral' bis 'sehr zufrieden': 89 % 19/126 Patient*innen unzufrieden mit der App, Erklärungen: Krankheit konnte nicht vollständig dargestellt werden: 62 %</p>

Autor*innen, Jahr, Quelle	Ronicke et al., 2019 [76]	Schmieding et al., 2021 [80]	Trivedi et al., 2021 [77]	Verzantvoort et al., 2018 [78]
Bedienbarkeit der digitalen Gesundheitsanwendung (Fortsetzung)				App-Empfehlungen nicht umfassend, zu schnelle Empfehlung: 24 % App strukturell mangelhaft: 14 %
Haltung der Patient*innen gegenüber Symptom-Checker	NA	NA	NA	<p>Absicht, App-Empfehlung zu befolgen: 65 %</p> <p>39/126 Patient*innen ohne Absicht, die App-Empfehlung zu befolgen: Krankheit konnte nicht vollständig dargestellt werden: 33 % Es fand bereits eine Ärzt*in-Konsultation statt: 27 % Mehr Vertrauen in das eigene Urteilsvermögen: 27 %</p> <p>Vorschläge zur Verbesserung der App: Bessere und vollständigere Erfassung aller Aspekte der Krankheit ermöglichen; Struktur, Bedienung und Geschwindigkeit der App verbessern; Empfehlungen der App verbessern, Layout der App verbessern</p>

Abkürzungen: CI – Konfidenzintervall; NA – nicht verfügbar (not available); NR – nicht berichtet (not reported); p – p-Wert; SD – Standardabweichung (standard deviation); UK – United Kingdom

Studien zur Patient*innensicht aus der Update-Suche

Tabelle A-4: Datenextraktionstabelle zu zusätzlich identifizierten Studien zur Patient*innensicht von Symptom-Checkern

Study name, author, year	Aboueid et al., 2021 [81]	Meyer et al., 2020 [83]	Miller et al., 2020 [84]	Eccles et al., 2019 [82]	Nadarzynski et al. 2019 [85]	Winn et al., 2019 [86]
Country	Canada	USA	UK/Germany	UK	UK	USA
Sponsor/ Funding	Funding not reported. No conflicts of interest declared.	Supported in part by the Houston Veterans Administration Health Services Research and Development; Agency for Healthcare Research and Quality Mentored Career Development Award; Gordon and Betty Moore Foundation	Funding not reported. Authors employees or paid consultants of Ada Health.	Bursary from University of Warwick's Undergraduate Research Support Scheme	No specific funding received	Supported by the National Center for Research Resources, the National Center for Advancing Translational Sciences, and the office of the Director, National Institutes of Health
Study design	Qualitative descriptive case study	Cross-sectional survey study	Exploratory pilot study	Mixed-methods retrospective exploratory study	Mixed-methods retrospective study	Cross-sectional study
Intervention	WebMD Symptom Checker; Babylon Health	Isabel Symptom Checker	Ada Health	askmyGP	Any health chatbot	Buoy Health
Comparator	Google search engine	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
„Approach“/ Setting	Semistructured interviews about experience using a symptom checker with a clinical vignette.	Survey distributed to users that used the Symptom checker within the last 6 months	Study questionnaire about app perception and app utility distributed to patients after completing the Ada self-assessment about their presenting complaint	Retrospective analysis of routinely collected data from patients, the data came from nine general practices in the UK	Semi-structured interviews and an online survey distributed on social media to explore motivations for the use of chatbots in healthcare	Data from patient encounters using Buoy Health were analyzed; chief concerns were categorized based on organ system, concern, and location. Preencounter intent (intended level of care at the beginning) and postencounter intention data were collected.
Number of users	24 young adults enrolled in a university in Ontario, Canada	385 respondents with 329 complete respondents	523 participants	5,447 patients	29 participants (qualitative sub-study) 215 participants (quantitative sub-study)	158,083 encounters
User characteristics	Age: 18-34 (no mean age reported) Women: 58% Men: 38% Nonbinary: 4%	Mean respondent age: 48,0 (SD 16,7) Women: 75,7% White: 89,1%	Mean age: 39,79 (SD 17,7) Women: 62,1%	Mean age: 36 (range 0-95) Women: 65,5%	Qualitative sub-study: Women: 52% White: 83% Quantitative sub-study: Mean age: 30 (SD 12) Women: 61% White: 64% Educated below university degree: 54% Self-rated good or very good IT skills: 76%	Mean age: 40 (SD 16,5) Women: 78% Reported symptoms: 281,165 Symptoms with identifiable organ systems: 204,958 Reproductive organ system: 18% General system: 17% Gastrointestinal system: 15%

Study name, author, year	Aboueid et al., 2021 [81]	Meyer et al., 2020 [83]	Miller et al., 2020 [84]	Eccles et al., 2019 [82]	Nadarzynski et al. 2019 [85]	Winn et al., 2019 [86]
User characteristics (continuation)						Symptoms with identifiable concern types: 279,948 Pain: 44% Abnormal functioning: 22% Discharge: 8%
Experience around symptom checker use	NA	<p>To better understand cause of symptoms: 76,3%</p> <p>To decide whether to seek in-person health care: 33,2%</p> <p>To decide what health care setting to visit: 20,7%</p> <p>To get medical advice without going to the doctor: 15,8%</p> <p>To better understand diagnosis made by their doctor: 12,8%</p> <p>Visiting the ED after receiving advice to proceed to ED: 54%</p> <p>Information of symptom checker considered useful: 90,1%</p> <p>Positive health effects reported: 51,0%</p> <p>Perceived as a useful diagnostic tool: 84,1%</p> <p>Perceived as a tool providing insights leading closer to correct diagnosis: 76,2%</p> <p>Reported to use symptom checker again: 91,4%.</p> <p>Discussed findings with physician: 48,4%</p> <p>Felt physician were interested in findings: 40,8%</p> <p>Felt physicians were not interested in findings: 23,3%.</p> <p>Felt physicians were open to discuss findings: 60,2%</p> <p>Felt physicians were not open to discuss findings: 20,4%</p> <p>Patients with previous experienced diagnostic errors:</p> <p>Likely to use symptom checker to determine where to seek care: 26,5%</p> <p>Felt physician interested in discussing findings: 32%</p>	<p>Use of symptom checker rated as very or quite easy: 97,8%</p> <p>Reported to use symptom checker again: 88,1%</p> <p>Reported to recommend symptom checker to friend or relative: 85,3%</p> <p>Reported helpful advice from symptom checker: 18-24 year-olds: 93% 70+ year olds: 59%</p> <p>Reported that use of symptom checker made no difference in care-seeking behaviour: 86,0%</p> <p>Reported that use of symptom checker resulted in seeking lower-intensity care: 12,8%</p> <p>18-24 year olds: 22%</p> <p>70+ year olds: 0%</p>	<p>Patients commented the need to use the platform out of necessity (due to busy phone lines, or by signposting by reception staff to use the platform in lieu of booking an appointment).</p> <p>General dissatisfaction with current routes for accessing general practitioners, including online triage.</p>	<p>Had heard of health chatbots: 6%</p> <p>Had experience using health chatbots: 3%</p> <p>Perceived utility:</p> <p>Likely to seek general health information: 78%</p> <p>Likely to seek information about medication: 60%</p> <p>Likely to seek information about diseases: 69%</p> <p>Likely to seek information about symptoms: 67%</p> <p>Likely to seek results of medical tests: 39%</p> <p>Likely to book a medical appointment: 78%</p> <p>Likely to look for local medical services: 80%</p> <p>Likely to seek specialist advice: 48%</p> <p>Beliefs associated with chatbot acceptability:</p> <p>Worried about health: 50%</p> <p>Worried about privacy using a health chatbot: 47%</p> <p>Worried about the security of information: 47%</p> <p>Confident in finding accurate health information online: 48%</p> <p>Confident in identifying own health symptoms: 68%</p> <p>Comfortable with outlining symptoms to a chatbot: 61%</p> <p>Prefer to talk face to face with a doctor about health: 73%</p> <p>Dislike of talking to computers: 30%</p> <p>Strange to talk to a chatbot about health: 41%</p> <p>Health chatbot could help to make better decisions: 30%</p> <p>Would trust advice from a health chatbot: 27%</p> <p>A health chatbot is a good idea: 65%</p>	<p>Preencounter intended level of care:</p> <p>Primary care: 46,52%</p> <p>Uncertain: 33,94%</p> <p>Urgent care: 9,44%</p> <p>Emergency care: 9,37%</p> <p>Telemedicine: 0,73%</p> <p>Postencounter intended level of care:</p> <p>Primary care: 39,57%</p> <p>Uncertain: 20,52%</p> <p>Wait and watch: 17,26%</p> <p>Self-treatment: 9,99%</p> <p>Urgent care: 5,93%</p> <p>Emergency care: 6,09%</p> <p>Telemedicine: 0,64%</p>

Study name, author, year	Aboueid et al., 2021 [81]	Meyer et al., 2020 [83]	Miller et al., 2020 [84]	Eccles et al., 2019 [82]	Nadarzynski et al. 2019 [85]	Winn et al., 2019 [86]
<p>Experience around symptom checker use <i>(continuation)</i></p>		<p><i>Patients without previous experienced diagnostic errors:</i> Likely to use symptom checker to determine where to seek care: 12,2% Felt physician interested in discussing findings: 59%</p>			<p>Willing to enter symptoms on an online form: 63% Curious how new technologies could improve health: 80% Reliable and accurate information is important: 93% Only seek a doctor if there is a health problem: 66%</p> <p>Understanding of chatbots: <i>Awareness:</i> „I think that it's online and you ask it questions and it can reply to you with information. It is not a real person. It is like stored information.” „There are those diagnosis bots and you can use. Those to get an idea of what is wrong with you or what your next step could be if it can just tell you a rough severity of what you may have. Then I suppose it is useful in those cases but, in terms of just chatting to something and telling it your problems or whatever and trying to get a diagnosis. If you are using it in that kind of way then I think that is a very limited thing to try and do.” „I've used one for banking before, and you just type in your query. And I think I've used some which you type to and it picks out keywords, and one which you can select the most appropriate response.” <i>Experience:</i> „Where it talks back to you, it can be more specific, compared to Google where you have to search and look through.”</p>	
<p>Positive themes</p>	<p>Symptom-Checker: More personalized; more interactive due to chatbot feature; good for those who do not know how to use Google; straightforward design; easy to use; real-time output; makes the correlation between symptoms and potential conditions; more intuitive; more reliable; more specific; more structured</p>	<p>NA</p>	<p>NA</p>	<p>Removal of the need for verbal interaction viewed positively by some; for certain health conditions non-verbal form of communication is easier Improved quality of communication as there was time to write about what was wanted to discuss.</p>	<p>Motivations for health chatbots: <i>Anonymity:</i> „I think for mental health it would be pretty useful because I think that it's a lot harder to talk to a real person about that. Maybe sexual health too. I'm pretty open generally about both of those things but I can see where they might be seen as a better alternative due to privacy and not having to face a person and describe all of these problems.”</p>	<p>NA</p>

Study name, author, year	Aboueid et al., 2021 [81]	Meyer et al., 2020 [83]	Miller et al., 2020 [84]	Eccles et al., 2019 [82]	Nadarzynski et al. 2019 [85]	Winn et al., 2019 [86]
Positive themes (continuation)	Google search engine: Provides information without claiming a diagnosis; more customizable; allows entry of all symptoms in the search engine			Some users found the platform particularly appropriate for certain problems. Users who were unsure whether the problem warranted an appointment considered the service as a useful way of checking. Some viewed the system as a potential way to save general practitioner's time and prioritise appointments. Some patients reported the system working better than booking appointments via the phone.	„I think mental health would be a good thing to use a chatbot in, because some people with mental health issues they do not want to open up to an actual person, so it would be easier doing it over the internet in the comfort of their own home.“ <i>Convenience:</i> „You can use a chatbot instead of googling it and reading advice on the NHS page. If that information is integrated into the chatbot then yeah it would certainly save time.“ „It would be good for healthcare because obviously, it is so hard to get a doctor's appointment, so for people who have general queries like it would be good for them to quickly get advice on whether or not it is an urgent issue.“ <i>Sign-posting:</i> „Chatbot can tell you if you need further advice or you will worry. And it will reassure you so you don't go to the doctors.“ „People who have no awareness of or are perhaps too worried, it could be a good way to get in touch. Just like if you use 111, first of all, you go to a chatbot and they would figure out what is wrong and if there was a severity to it they would ring 999 and get an ambulance, otherwise, they can direct you to a GP and a clinic.“	
Negative themes	Symptom checkers: Accuracy is questionable; limits the number of symptoms that can be inputted; not widely known; thought process of the platform is unclear; user more vulnerable when using this platform Google search engine: Absence of chatbot feature Both symptom checkers and Google search engine: Text input is insufficient; suboptimal reliability			Lack of verbal interaction; not wanting to explain to a computer as opposed to a person For patients with sight problems the online visual element is a disadvantage Dislike of asynchronous interaction offered by the platform Concerns whether patients can be adequately assessed remotely.	Hesitancy: <i>Perceived accuracy:</i> „It's not that I think they [chatbots] would intentionally give me false information. I just don't know how accurate they are. I don't know whether they can be as accurate as doctors can be.“ „I would find it hard to trust a health chatbot because it is just looking online at things. You would want a professional opinion.“	NA

Study name, author, year	Aboueid et al., 2021 [81]	Meyer et al., 2020 [83]	Miller et al., 2020 [84]	Eccles et al., 2019 [82]	Nadarzynski et al. 2019 [85]	Winn et al., 2019 [86]
<p>Negative themes (continuation)</p>		<p>NA</p>	<p>NA</p>	<p>Difficult logistical process of taking the telephone call response from the practice after the initial online encounter. Some patients reported missing or not receiving the follow-up phone call.</p> <p>Concerns about the quality of the description in reference to the patient's problem.</p> <p>The platform considered less appropriate when the patient was not sure what was wrong.</p> <p>When the patient wanted to book a face-to-face appointment, the platform was considered as a barrier.</p> <p>Concern about the need for patients to have computer skills. Some computer-savvy patients expressed this concern, even though they were not directly invited to comment on suitability for others.</p>	<p><i>AI Premature technology:</i> „I think it has a lot of potential in the future, but now certain places are releasing it before they have perfected their own system, then it could put people off. Because you can end up chatting for like half an hour and go back to being at the same question you were at in the first place. With that people get angry.”</p> <p>„I don't know whether the technology is as adequate as a doctor.”</p> <p><i>Non-human interaction:</i> „I think a lot of people would be put off just being with a chatbot. It is like a segregation thing, I don't think it will replace human interaction.”</p> <p>„If you are looking at a chatbot thinking that it is a replacement for a person then you are looking it in the wrong way. If you were looking for a deep and meaningful conversation you are not going to find one.”</p> <p><i>Cyber-security:</i> „Some people might find issues with confidentiality because if you were with the doctor it is just you and them, but with chatbots, you don't know who is behind it all.”</p> <p>„Some things are confidential and you wouldn't just type it on the internet. You would want the confidentiality of the GP practice.”</p>	
<p>Enablers</p>	<p><i>Individual-level factors:</i> Internet access; low health literacy; trust in the platform; young age; lack of time; convenience; lack of trust in doctors; curiosity; embarrassing topic; increase empowerment; aversion to medical professionals; having pre-existing conditions; unable to discuss the topic with a health provider; uncertain about care required; worried about health of oneself</p> <p><i>Disease-level factors</i> Mild symptoms; a „broad category of illness”; symptoms can be easily described</p>	<p>NA</p>	<p>NA</p>	<p>NA</p>	<p>NA</p>	<p>NA</p>

Study name, author, year	Aboueid et al., 2021 [81]	Meyer et al., 2020 [83]	Miller et al., 2020 [84]	Eccles et al., 2019 [82]	Nadarzynski et al. 2019 [85]	Winn et al., 2019 [86]
Enablers (<i>continuation</i>)	<p>Health system-level factors: Approved by doctors; lack of access to health services; cost of health services; public education; increased awareness; long wait times for health services; reputable organizations recommend it</p> <p>Characteristics of symptom checkers: Increased advertisement; easy interface; data privacy; free of charge; short to complete; precision; use of artificial intelligence; gamification; good source of information; integrated with an electronic health record; useful in identifying potential conditions; information about the creators of the platform; interactive platform; reliability</p>					
Barriers	<p>Individual-level factors: Lack of internet access; low health literacy; lack of trust in the platform; low technology literacy; older age; social influence; not wanting to know; previous bad experience</p> <p>Disease-level factors: Severe condition; need for a physical examination</p> <p>Health system-level factors: Authoritarianism in health care</p> <p>Characteristics of symptom checkers: Lack of awareness; poor design; asking identifiable questions; cost of the platform; time to complete; lack of inclusivity measures; lack of language options; lack of credibility; lack of human interaction; disclaimer; inability to obtain elaboration on a question; liability; concerns about using data for profits</p>	NA	NA	NA	NA	NA

Abkürzungen; ED: Notaufnahme („emergency department“); GP: Allgemeinmediziner*in („General practitioner“); NA: nicht verfügbar („not available“); SD: Standardabweichung („standard deviation“); UK: United Kingdom; USA: United States of America

Risk-of-Bias Assessment für diagnostische Studien mit dem QUADAS-II Assessment Tool

Tabelle A-5: Einstufung des Verzerrungspotenzials (Risk-of-Bias) für diagnostische Studien unter Anwendung des QUADAS-II Bewertungsinstruments [28]

Autor*innen, Jahr, Quelle	Risk of Bias				Applicability concerns		
	Patient selection (Domain 1)	Index test (Domain 2)	Reference standard (Domain 3)	Flow and timing (Domain 4)	Patient selection (Domain 1)	Index test (Domain 2)	Reference standard (Domain 3)
Baker et al., 2020 [73]	H No random patient sample enrolled (Patients were played by general practitioners, of whom some were employees or consultants of Babylon Health)	L Unclear if index test results were interpreted without knowledge of the results of the reference standard (Possibility that algorithm was trained with the clinical vignettes)	H Doctors assessing the clinical vignettes were used as reference standards; some of the doctors provide consultations for Babylon Health. Unclear if knowledge of results of the index test were present	H Not all clinical vignettes received the same reference standard	H Study population does not represent the average target population	L	H Not all possible symptoms covered by clinical vignettes
Nateqi et al., 2019 [74]	H No real patients were enrolled	U Unclear if index test results were interpreted without knowledge of the results of the reference standard (Possibility that algorithm was trained with the clinical vignettes)	H No adequate definition of reference standard	L	H Study population doesn't represent the average target population	L	H Not all possible symptoms covered by clinical vignettes
Gilbert et al., 2020 [75]	H High risk of bias in patient selection, no random sample enrolled (Some of the doctors assessing the clinical vignettes were previous employees of Ada Health)	U Unclear if index test results were interpreted without knowledge of the results of the reference standard (Possibility that algorithm was trained with the clinical vignettes)	H Employees of Ada Health were involved in the creation of the clinical vignettes Unclear if the reference standard had knowledge of index test results	L	H Study population does not represent the average target population (The entry of symptoms was conducted by healthcare professionals instead of laypeople)	L	H Not all possible symptoms covered by clinical vignettes
Ronicke et al., 2019 [76]	H No consecutive or random enrollments of patients (Selection was based on subjective evaluation by a clinician)	U Unclear if index test results were interpreted without knowledge of the results of the reference standard	L	L	H Study population doesn't represent the average target population (Setting: clinic for rare inflammatory systemic diseases)	L	H Not all possible symptoms covered by selected clinical cases

Autor*innen, Jahr, Quelle	Risk of Bias				Applicability concerns		
	Patient selection (Domain 1)	Index test (Domain 2)	Reference standard (Domain 3)	Flow and timing (Domain 4)	Patient selection (Domain 1)	Index test (Domain 2)	Reference standard (Domain 3)
Trivedi et al., 2021 [77]	H No random patient sample enrolled (The study population was a convenience sample of patients arriving at the emergency department without an ambulance)	L	U Unclear if reference standard is likely to correctly classify the target condition	L	L	L	U Unclear if the target condition as defined by the reference standard matches the review question?
Verzantvoort et al., 2018 [78]	L	L	L	L	L	L	L
Hill et al., 2020 [79]	H No real patients were enrolled	U Unclear if index test results were interpreted without knowledge of the results of the reference standard (Possibility that algorithm was trained with the clinical vignettes)	U Unclear if reference standard is likely to correctly classify the target condition	L	H Study population doesn't represent the average target population	L	H Not all possible symptoms covered by clinical vignettes
Schmieding et al., 2021 [80]	H No real patients were enrolled	U Unclear if index test results were interpreted without knowledge of the results of the reference standard (Possibility that algorithm was trained with the clinical vignettes)	U Unclear if reference standard is likely to correctly classify the target condition	L	H Study population doesn't represent the average target population	L	H Not all possible symptoms covered by clinical vignettes

Legende: H: Hoch („High“); L: Niedrig („Low“); U: Unklar („Unclear“);

Meta-Verzeichnis von digitalen Gesundheitsanwendungen

Tabelle A-6: Meta-Verzeichnis der deutschsprachig verfügbaren digitalen Gesundheitsanwendungen aus ausgewählten Verzeichnissen (kategorisiert nach Zielgruppe und Funktion).
 Quellen der Verzeichnisse: [5, 91-96]

Name der App (Hersteller)	Plattformen	Verzeichnis	Anwender	Zielgruppe	Funktion der DiGA
EpiCure	App (Google Playstore)	AT – AIT	Patient*innen	0 Gesund ohne bekannte Risikofaktoren	0 Dokumentation
Medloop	App (Google Playstore & Apple Appstore)	UK – NHS Apps Library	Patient*innen und Kliniker*innen	0 Gesund ohne bekannte Risikofaktoren	0 Dokumentation
Bildschirmarbeit	App (Google Playstore & Apple Appstore)	AT – IfGP – Gesundheits-Apps	Patient*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
Bingli (Bingli)	Web and App	BE – mHealthBelgium	Patient*innen und Kliniker*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
Brush DJ	App (Google Playstore & Apple Appstore)	UK – NHS Apps Library	Patient*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
EKO2gO	Web and App	AT – IfGP – Gesundheits-Apps	Patient*innen und Kliniker*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
eQuoo: Emotional Fitness Game	App (Google Playstore & Apple Appstore)	UK – NHS Apps Library	Patient*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
Erste Hilfe Hand	App (Google Playstore & Apple Appstore)	AT – IfGP – Gesundheits-Apps	Patient*innen und Kliniker*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
Evergreen Life	Web and App	UK – NHS Apps Library	Patient*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
Food Jungle	App (Google Playstore & Apple Appstore)	AT – IfGP – Gesundheits-Apps	Patient*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
Heben und Tragen	App (Google Playstore & Apple Appstore)	AT – IfGP – Gesundheits-Apps	Patient*innen und Kliniker*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
Lexikon Prävention	Web and App	AT – IfGP – Gesundheits-Apps	Patient*innen und Kliniker*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
MedBusters	App (Google Playstore & Apple Appstore)	AT – IfGP – Gesundheits-Apps	Patient*innen und Kliniker*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
PACSonWEB (Dobco Medical Systems)	Web and App	BE – mHealthBelgium	Patient*innen und Kliniker*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
Sicherheits- und Gesundheitsmanagement (SGM)	Web and App	AT – IfGP – Gesundheits-Apps	Patient*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
Unfallfrei von Anfang an	App (Google Playstore & Apple Appstore)	AT – IfGP – Gesundheits-Apps	Patient*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information

Name der App (Hersteller)	Plattformen	Verzeichnis	Anwender	Zielgruppe	Funktion der DiGA
Well@Home	Web and App	BE – mHealthBelgium	Patient*innen und Kliniker*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
wellabe	k.A.	AT – HealthHubVienna Alumni	Patient*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
Zeichen der Sicherheit	App (Google Playstore & Apple Appstore)	AT – IfGP – Gesundheits-Apps	Patient*innen	0 Gesund ohne bekannte Risikofaktoren	1 Information
...von Anfang an!	App (Google Playstore & Apple Appstore)	AT – IfGP – Gesundheits-Apps	Patient*innen	1 Gesund mit Risikofaktoren	1 Information
Rauchfrei-App	App (Google Playstore & Apple Appstore)	AT – IfGP – Gesundheits-Apps	Patient*innen	1 Gesund mit Risikofaktoren	1 Information
SyncVR Fit (SyncVR Medical B.V.)	Webanwendung	BE – mHealthBelgium	Patient*innen	2 akut erkrankt, nicht lebensbedrohend	1 Information
SyncVR Relax & Distract (SyncVR Medical B.V.)	Webanwendung	BE – mHealthBelgium	Patient*innen	2 akut erkrankt, nicht lebensbedrohend	1 Information
Nucleus Smart App (Cochlear)	App (Google Playstore & Apple Appstore)	BE – mHealthBelgium	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	1 Information
image biopsy lab	Webanwendung	AT – HealthHubVienna Alumni	Kliniker/Gesundheitspersonal und Gesundheitsinstitutionen	0 Gesund ohne bekannte Risikofaktoren	4 Diagnosestellung
infermedica	Web and App	AT – HealthHubVienna Alumni	Patient*innen und Kliniker*innen	0 Gesund ohne bekannte Risikofaktoren	4 Diagnosestellung
Microsonix	k.A.	AT – HealthHubVienna Alumni	Kliniker/Gesundheitspersonal und Gesundheitsinstitutionen	0 Gesund ohne bekannte Risikofaktoren	4 Diagnosestellung
Symptoma – better diagnosis	Webanwendung	AT – HealthHubVienna Alumni	Patient*innen und Kliniker*innen	0 Gesund ohne bekannte Risikofaktoren	4 Diagnosestellung
Skinvision (SkinVision BV)	Web and App	BE – mHealthBelgium	Patient*innen	1 Gesund mit Risikofaktoren	4 Diagnosestellung
Sunrise (Sunrise)	App (Google Playstore & Apple Appstore)	BE – mHealthBelgium	Patient*innen und Kliniker*innen	1 Gesund mit Risikofaktoren	4 Diagnosestellung
Homemonitoring	App (Google Playstore)	AT – AIT	Patient*innen und Kliniker*innen	0 Gesund ohne bekannte Risikofaktoren	5 Einfaches Monitoring
Thryve	App (Google Playstore & Apple Appstore)	AT – HealthHubVienna Alumni	Patient*innen	0 Gesund ohne bekannte Risikofaktoren	5 Einfaches Monitoring
oscar senior	App (Google Playstore & Apple Appstore)	AT – HealthHubVienna Alumni	Patient*innen	1 Gesund mit Risikofaktoren	5 Einfaches Monitoring
Florio (Swedish Orphan Biovitrum (Belgium) BV)	Web and App	BE – mHealthBelgium	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	5 Einfaches Monitoring
icompanion (icomatrix)	Web and App	BE – mHealthBelgium	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	5 Einfaches Monitoring

Name der App (Hersteller)	Plattformen	Verzeichnis	Anwender	Zielgruppe	Funktion der DiGA
Noona (Varian Medical Systems Belgium)	App (Google Playstore & Apple Appstore)	BE – mHealthBelgium	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	5 Einfaches Monitoring
CardiacSense (Arseus Hospital NV)	Webanwendung	BE – mHealthBelgium	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring
Comarch Home Health (Comarch)	App (Google Playstore & Apple Appstore)	BE – mHealthBelgium	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring
Comunicare (Comunicare Solutions)	App (Google Playstore & Apple Appstore)	BE – mHealthBelgium	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring
FibriCheck (Qompium)	Web and App	BE – mHealthBelgium	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring
Freestyle LibreLink – BE (Abbott SA/NV)	App (Google Playstore & Apple Appstore)	BE – mHealthBelgium	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring
Guardian Connect App (Medtronic Belgium)	App (Google Playstore & Apple Appstore)	BE – mHealthBelgium	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring
HumanITcare	App (Google Playstore & Apple Appstore)	AT – ScaleUp4Europe	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring
MobECG (Savvy)	App (Google Playstore)	AT – ScaleUp4Europe	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring
Plateforme Maela (Medtronic Belgium)	Web and App	BE – mHealthBelgium	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	6 Komplexes Monitoring
AirviewTM (Resmed)	App (Google Playstore & Apple Appstore)	BE – mHealthBelgium	Kliniker/Gesundheitspersonal und Gesundheitsinstitutionen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
CANKADO PRO-React Onco (Cankado Service GmbH)	Web and App	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
CardioMemory	App (Google Playstore)	AT – AIT, IfGP – Gesundheits-Apps	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
Catch It	App (Google Playstore & Apple Appstore)	UK – NHS Apps Library	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
Deprexis (GAIA AG)	Webanwendung	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
DiabCare	App (Google Playstore)	AT – AIT	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
DiabMemory 2	App (Google Playstore & Apple Appstore)	AT – AIT	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
Elevida (GAIA AG)	Webanwendung	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
Heartfish	App (Google Playstore & Apple Appstore)	AT – ScaleUp4Europe	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention

Name der App (Hersteller)	Plattformen	Verzeichnis	Anwender	Zielgruppe	Funktion der DiGA
HerzMobil	App (Google Playstore)	AT – AIT, IfGP – Gesundheits-Apps	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
Mika (Fosanis GmbH)	App (Google Playstore & Apple Appstore)	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
mindmate	Web and App	AT – HealthHubVienna Alumni	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
M-sense Migräne (Newsenselab GmbH)	App (Google Playstore & Apple Appstore)	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
mySugr Tagebuch-App	App (Google Playstore & Apple Appstore)	AT, BE, UK – IfGP – Gesundheits-Apps; mHealthBelgium; UK – NHS Apps Library	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
Rehappy (Rehappy GmbH)	App (Google Playstore & Apple Appstore)	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
Selfapys Online-Kurs bei Depressionen (Selfapy GmbH)	Webanwendung	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
somnio (mementor DE GmbH)	Web and App	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
velibra (GAIA AG)	Webanwendung	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
Vitadio	App (Google Playstore & Apple Appstore)	AT – ScaleUp4Europe	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
Vivira (Vivira Health Lab GmbH)	App (Google Playstore & Apple Appstore)	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
vorvida (GAIA AG)	Webanwendung	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
zanadio (aidhere GmbH)	App (Google Playstore & Apple Appstore)	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
vaica	Web and App	AT – HealthHubVienna Alumni	Patient*innen und Kliniker*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	7 Indirekte Intervention
Invirto (Sympatient GmbH)	App (Google Playstore & Apple Appstore)	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	8 Direkte Intervention
Kalmeda (mynoise GmbH)	App (Google Playstore & Apple Appstore)	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	8 Direkte Intervention
Mindable: Panikstörungen und Agoraphobie (Mindable Health GmbH)	App (Google Playstore & Apple Appstore)	DE – DiGA-Verzeichnis	Patient*innen	3 chronisch erkrankt, stabiler Gesundheitszustand	8 Direkte Intervention

Legende: AT – Österreich; BE – Belgien; DE – Deutschland; UK – Großbritannien

Abkürzungen: AIT: Austrian Institute of Technology; DiGA: digitale Gesundheitsanwendung; IfGP: Institut für Gesundheitsförderung und Prävention

Suchstrategien

Suchstrategie für Cochrane Datenbank

Search date: 20/05/2021	
ID	Search
#1	MeSH descriptor: [Symptom Assessment] explode all trees
#2	(self-diagnos*) (Word variations have been searched)
#3	#1 OR #2 (Word variations have been searched)
#4	MeSH descriptor: [Software] explode all trees
#5	MeSH descriptor: [Artificial Intelligence] explode all trees
#6	MeSH descriptor: [Telemedicine] explode all trees
#7	#4 OR #5 OR #6 (Word variations have been searched)
#8	#3 AND #7 (Word variations have been searched)
#9	(„symptom* checker*“) (Word variations have been searched)
#10	(„self*-triag*“) (Word variations have been searched)
#11	(web-based NEAR triag*) (Word variations have been searched)
#12	(„digital health assistan*“) (Word variations have been searched)
#13	(„health* chatbot*“) (Word variations have been searched)
#14	(„Diagnostic decision support system*“) (Word variations have been searched)
#15	(symptom* NEAR app?):ti,ab,kw
#16	#8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15
Total: 199 Hits	

Suchstrategie für Embase

No.	Query Results	Results	Date
#53.	#52 AND ([english]/lim OR [german]/lim)	697	20 May 2021
#52.	#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42 OR #43 OR #44 OR #45 OR #46 OR #47 OR #48 OR #49 OR #50 OR #51	707	20 May 2021
#51.	webgp	1	20 May 2021
#50.	webmd:dn,df	4	20 May 2021
#49.	symptomate	3	20 May 2021
#48.	symptoma AND (web OR internet OR online OR app\$)	2	20 May 2021
#47.	symptify	-	20 May 2021
#46.	symcat	2	20 May 2021
#45.	steps2care	-	20 May 2021
#44.	'push doctor'	2	20 May 2021
#43.	'nhs symptom checker*'	1	20 May 2021
#42.	mediktor	2	20 May 2021
#41.	'me\$doctor'	-	20 May 2021
#40.	itriage	7	20 May 2021
#39.	isabel:dn	7	20 May 2021
#38.	healthychildren.org	10	20 May 2021
#37.	(health\$line OR health\$wise OR 'mayo clinic':dn) AND (symptom*:ti,ab OR diagnos*:ti,ab)	50	20 May 2021
#36.	health\$direct	24	20 May 2021
#35.	'harvard medical school family health guide'	1	20 May 2021
#34.	'gp at hand'	18	20 May 2021

#33.	free\$md	2	20 May 2021
#32.	find\$zebra	10	20 May 2021
#31.	'family\$doctor'	5	20 May 2021
#30.	esagil	-	20 May 2021
#29.	'engage consult'	-	20 May 2021
#28.	econsult:dn	1	20 May 2021
#27.	'early doc'	4	20 May 2021
#26.	'early\$doc'	-	20 May 2021
#25.	'earlydoc':dn	-	20 May 2021
#24.	'drugs.com':dn	1	20 May 2021
#23.	doc\$response	-	20 May 2021
#22.	better\$medicine	-	20 May 2021
#21.	babylon:dn	3	20 May 2021
#20.	'ask md'	2	20 May 2021
#19.	ada:dn	8	20 May 2021
#18.	'online triage'	18	20 May 2021
#17.	(diagnos* OR symptom*) NEAR/1 app\$	96	20 May 2021
#16.	'diagnostic decision support system*'	108	20 May 2021
#15.	'health* chatbot*'	17	20 May 2021
#14.	'digital health assistan*'	1	20 May 2021
#13.	'web-based triag*'	16	20 May 2021
#12.	'self*-triag*'	77	20 May 2021
#11.	'symptom* checker*'	90	20 May 2021
#10.	#3 AND #9	156	20 May 2021
#9.	#4 OR #5 OR #6 OR #7 OR #8	676,513	20 May 2021
#8.	'telemedicine'/exp	49,494	20 May 2021
#7.	'machine learning'/exp	250,197	20 May 2021
#6.	'artificial intelligence'/exp	47,634	20 May 2021
#5.	'software'/exp	387,992	20 May 2021
#4.	'mobile application'/exp	15,687	20 May 2021
#3.	#1 OR #2	2,405	20 May 2021
#2.	'self diagnos*'	1,115	20 May 2021
#1.	'symptom assessment'/mj	1,291	20 May 2021

Suchstrategie für Medline

Database: Ovid MEDLINE(R) and In-Process, In-Data-Review & Other Non-Indexed Citations and Daily <1946 to May 18, 2021>, Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily <2017 to May 18, 2021>	
Search Strategy:	
1	exp Symptom Assessment/(9598)
2	exp Software/(198867)
3	exp Artificial Intelligence/(155709)
4	exp Telemedicine/(47754)
5	2 or 3 or 4 (380280)
6	1 and 5 (231)
7	symptom* checker*.mp. (127)
8	self*-triag*.mp. (83)
9	web-based triag*.mp. (17)
10	digital health assistan*.mp. (2)

11	health* chatbot*.mp. (32)
12	Diagnostic decision support system*.mp. (113)
13	((diagnos* or symptom*) adj app\$1).mp. (60)
14	6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 (646)
15	limit 14 to (english or german) (635)
16	remove duplicates from 15 (396)
Search date: 19/05/2021	

Suchstrategie für CRD

#### Symptom Checkers (RJ/GG) 200521	
Search date: 20/05/2021	
1	MeSH DESCRIPTOR Symptom Assessment EXPLODE ALL TREES
2	(self-diagnos*)
3	#1 OR #2
4	MeSH DESCRIPTOR Software EXPLODE ALL TREES
5	MeSH DESCRIPTOR Artificial Intelligence EXPLODE ALL TREES
6	MeSH DESCRIPTOR Telemedicine EXPLODE ALL TREES
7	#4 OR #5 OR #6
8	#3 AND #7
9	(symptom* checker*)
10	(self*-triag*)
11	(web-based triag*)
12	(digital health assistan*)
13	(health* chatbot*)
14	(Diagnostic decision support system*)
15	((diagnos* OR symptom*) NEAR app)
16	((diagnos* OR symptom*) NEAR apps)
17	#8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16
Total: 2 Hits	

Suchstrategie für INAHTA

Search Step: Search query,"Hits","Searched At"	
#11	((diagnos* OR symptom*) AND apps) OR ((diagnos* OR symptom*) AND app) OR („Diagnostic decision support system*“) OR (health* chatbot*) OR („digital health assistan*“) OR („web-based triag*“) OR („self*-triag*“) OR (symptom* checker*) OR („self-diagnos*“) OR („Symptom Assessment“[mhe]),“7“,“2021-05-20T17:28:46.000000Z“
#10	(diagnos* OR symptom*) AND apps,“2“,“2021-05-20T17:27:41.000000Z“
#9	(diagnos* OR symptom*) AND app,“2“,“2021-05-20T17:27:30.000000Z“
#8	„Diagnostic decision support system*“,“0“,“2021-05-20T17:26:59.000000Z“
#7	health* chatbot*,“0“,“2021-05-20T17:26:26.000000Z“
#6	„digital health assistan*“,“0“,“2021-05-20T17:26:05.000000Z“
#5	„web-based triag*“,“0“,“2021-05-20T17:25:11.000000Z“
#4	„self*-triag*“,“0“,“2021-05-20T17:24:08.000000Z“
#3	symptom* checker*,“2“,“2021-05-20T17:23:26.000000Z“
#2	„self-diagnos*“,“0“,“2021-05-20T17:22:01.000000Z“
#1	„Symptom Assessment“[mhe],“3“,“2021-05-20T17:14:18.000000Z“
Date of search: 20.05.2021	
5 Hits (#11 limited to English, German)	



HTA Austria
Austrian Institute for
Health Technology Assessment
GmbH