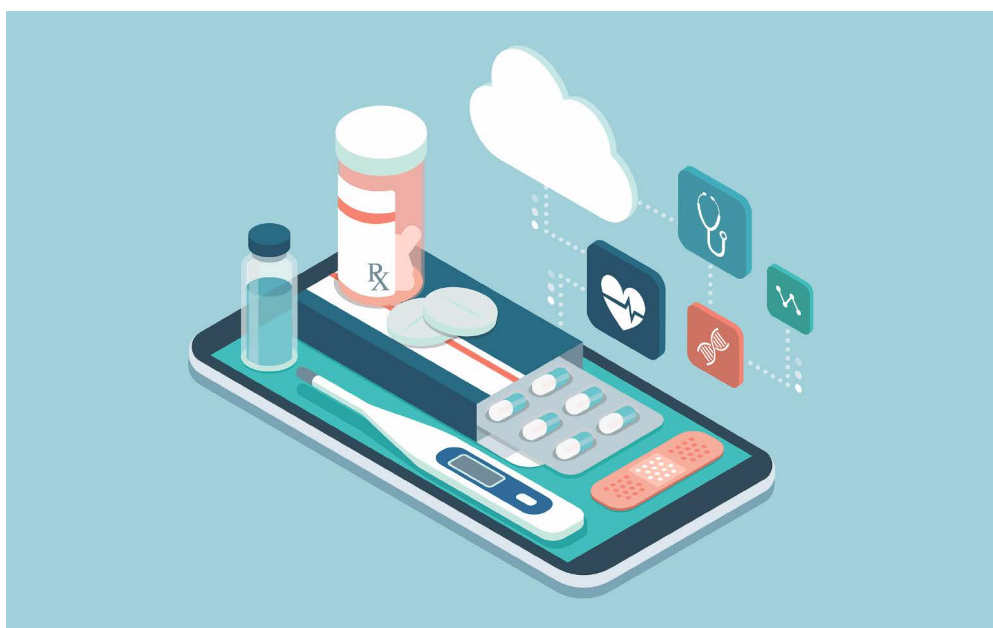


Framework zur Unterstützung von Refundierungsentscheidungen zu digitalen Gesundheitsanwendungen



... und dessen (retrospektive) Anwendung
an ausgewählten Beispielen



HTA Austria

Austrian Institute for
Health Technology Assessment
GmbH

Framework zur Unterstützung von Refundierungsentscheidungen zu digitalen Gesundheitsanwendungen

... und dessen (retrospektive) Anwendung
an ausgewählten Beispielen

Projektteam

Projektleitung: Priv. Doz. Dr.phil. Claudia Wild
Projektbearbeitung: Dr. med. Reinhard Jeindl
Priv. Doz. Dr.phil. Claudia Wild

Projektbeteiligung

Systematische Literatursuche: Tarquin Mittermayr, BA(Hons), MA
Evt. Handsuche: Kathrin Wohlhöfner
Externe Begutachtung: Dr. Dipl-Biol. Petra Schnell-Inderst, MPH, Institute of Public Health, Medical Decision Making and Health Technology Assessment, UMIT, Tirol
Interne Begutachtung: Dr. rer. soc. oec. Ingrid Zechmeister-Koss, MA

Korrespondenz: reinhard.jeindl@aihta.at

Umschlagfoto: © elenabsi – stock.adobe.com

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

Jeindl R, Wild C. Framework zur Unterstützung von Refundierungsentscheidungen zu digitalen Gesundheitsanwendungen (mHealth) und dessen (retrospektive) Anwendung an ausgewählten Beispielen. AIHTA Projektbericht Nr.: 134; 2020. 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.

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.: 134

ISSN 1993-0488

ISSN online 1993-0496

© 2020 AIHTA – Alle Rechte vorbehalten

Inhalt

Zusammenfassung	11
Executive Summary	13
1 Hintergrund.....	15
1.1 Einleitung.....	15
1.2 Definitionen.....	17
1.3 Zulassung und Risikoklassifizierung von digitalen Gesundheitsanwendungen	19
2 Methodisches Vorgehen.....	23
2.1 Projektziel	23
2.2 Forschungsfragen	23
2.3 Systematische Literatursuche.....	23
2.3.1 Ein- und Ausschlusskriterien.....	24
2.3.2 Literaturauswahl	24
2.4 INAHTA ListServ zu Assessment Frameworks	25
2.5 Auswahl und Datenextraktion der Assessment Frameworks	26
2.6 Identifikation von digitalen Gesundheitsanwendungen	26
2.6.1 Handsuche.....	26
2.6.2 Ein- und Ausschlusskriterien der digitalen Gesundheitsanwendungen	27
2.7 Auswahl der digitalen Gesundheitsanwendungen	28
2.8 Anwendung der Assessment Frameworks	28
2.9 Auswahl der Länder	28
3 Ergebnisse.....	29
3.1 Übersicht Assessment Frameworks.....	29
3.1.1 HTA Module for Mobile Medical Applications (MMAs) (AU).....	30
3.1.2 Evidence Standards Framework for Digital Health Technologies (ENG).....	31
3.1.3 Digi-HTA (FIN)	32
3.1.4 Digital Health Interventions (DEDHI) Framework (CH)	32
3.1.5 Digital Health Scorecard (USA)	33
3.1.6 mHealth Agile Development & Lifecycle (CA)	34
3.1.7 Weitere Frameworks	34
3.1.8 Zusammenfassung und Diskussion der ausgewählten Frameworks	38
3.2 Refundierte digitale Gesundheitsanwendungen in ausgewählten Ländern	40
3.2.1 Übersicht zu und Kurzbeschreibung von ausgewählten digitalen Gesundheitsanwendungen	40
3.2.2 Zusammenfassung und Diskussion der ausgewählten digitalen Gesundheitsanwendungen	44
3.3 Evidenz der ausgewählten digitalen Gesundheitsanwendungen	48
3.3.1 Verfügbare Evidenz zu den ausgewählten digitalen Gesundheitsanwendungen	48
3.3.2 Zusammenfassung und Diskussion der vorliegenden Evidenz zu ausgewählten digitalen Gesundheitsanwendungen	58
3.4 Anwendung der Assessment Frameworks	60
3.4.1 Anwendbarkeit der Frameworks	60
3.4.2 Zuteilung der ausgewählten digitalen Gesundheitsanwendung in Risikoklassen	61
3.4.3 Anwendung der HTA-Assessment Frameworks.....	63
3.4.4 Zusammenfassung und Diskussion der Anwendung der Assessment Frameworks	66
3.5 Übersicht von Regelungen in ausgewählten Ländern	68
3.5.1 Belgien	68
3.5.2 Frankreich	70
3.5.3 Niederlande	71
3.5.4 England.....	73
3.5.5 Deutschland	74
3.5.6 Österreich	77

4	Diskussion	81
4.1	Zusammenfassung: Kritische Bewertung der Frameworks und ihrer Anwendbarkeit	81
4.2	Weitere Aspekte und Bereiche mit besonderer Aufmerksamkeit.....	83
4.2.1	Datenschutz.....	83
4.2.2	Künstliche Intelligenz.....	84
4.2.3	Digital Health Literacy.....	85
4.2.4	Neue Studiendesigns.....	85
4.3	Limitationen der Arbeit	86
4.4	Schlussfolgerungen.....	87
5	Empfehlung für Framework für Österreich	89
6	Literatur.....	91
7	Anhang.....	99
7.1	mHealth Assessment Frameworks	99
7.1.1	HTA Module for Mobile Medical Applications (MMAs).....	99
7.1.2	Evidence Standards Framework for Digital Health Technologies.....	102
7.1.3	Digi-HTA Framework	104
7.1.4	Digital Health Interventions (DEDHI) Framework.....	106
7.1.5	Digital Health Scorecard	107
7.1.6	mHealth Agile Development & Lifecycle.....	108
7.2	Extraktionstabellen	109
7.2.1	INAHTA ListServ Antworten	109
7.3	Suchstrategien.....	112

Abbildungsverzeichnis

Abbildung 1.1-1:	Durchschnittliche Tagesnutzung digitaler Technologien in den USA 2008-2018	15
Abbildung 1.2-1:	WHO Classification of digital health interventions v1.0.....	18
Abbildung 2.3-1:	Darstellung des Auswahlprozesses (PRISMA Flow Diagram)	25
Abbildung 3.5-1:	Belgische Validierungspyramide	69
Abbildung 3.5-2:	Übersicht zum Ablauf des Fast-Track-Verfahrens.....	75
Abbildung 3.5-3:	Timeline der Umsetzung des Fast-Track-Verfahrens.....	77
Figure 7-1:	DHTs classified by function and stratified into evidence tiers	102
Figure 7-2:	NICE Framework Anforderungen an neue digitale Gesundheitsanwendungen, Visualisierung nach IQWiG	103
Figure 7-3:	Assessment process for digital healthcare services	105
Figure 7-4:	Components of Digital Health Scorecard. The four domains of a digital health scorecard with example considerations are detailed in this figure.....	107
Figure 7-5:	mHealth Agile Development & Evaluation Lifecycle	108

Tabellenverzeichnis

Tabelle 1.3-1:	Risikoklassen nach Signifikanz der Information für Entscheidungen, MDCG–Guidance	20
Tabelle 3.1-1:	Ausgewählte Frameworks zur Evaluation von mHealth Anwendungen	36
Tabelle 3.2-1:	Digitale Gesundheitsanwendungen – Anbieter, Funktionen und Zielgruppe	43
Tabelle 3.2-2:	Digitale Gesundheitsanwendungen – Versicherungen, Länder, Preis-/Refundierungsmodell	47
Tabelle 3.3-1:	Publizierte Studien zu SkinVision (CE-Kennzeichnung: 2014; derzeit geltende Risikoklasse nach MDD laut Hersteller: k.A.)	48
Tabelle 3.3-2:	Publizierte Studien zu ADA (CE-Kennzeichnung: 2016; derzeit geltende Risikoklasse nach MDD laut Hersteller: 1)	49

Tabelle 3.3-3:	Laufende Studien zu Mindable (CE-Kennzeichnung: vorhanden, keine Angabe zum Zeitpunkt; derzeit geltende Risikoklasse nach MDD laut Hersteller: 1)	50
Tabelle 3.3-4:	Publizierte Studien zu Natural Cycles (CE-Kennzeichnung: 2017; derzeit geltende Risikoklasse nach MDD laut Hersteller: 2b).....	50
Tabelle 3.3-5:	Publizierte Studien zu Preventicus Heartbeat (CE-Kennzeichnung: vorhanden, keine Angabe zum Zeitpunkt; derzeit geltende Risikoklasse nach MDD laut Hersteller: 2a)	52
Tabelle 3.3-6:	Publizierte Studien zu DiabiLive (CE-Kennzeichnung: 2019; derzeit geltende Risikoklasse nach MDD laut Hersteller: 2b).....	53
Tabelle 3.3-7:	Publizierte und aktuell laufende Studien zu Kaia (CE-Kennzeichnung: vorhanden, keine Angabe zum Zeitpunkt; derzeit geltende Risikoklasse nach MDD laut Hersteller: 1)	53
Tabelle 3.3-8:	Publizierte Studien zu Moodpath (CE-Kennzeichnung: 2016; derzeit geltende Risikoklasse nach MDD laut Hersteller: 1)	55
Tabelle 3.3-9:	Publizierte Studien zu Tinnitracks/TMNMt (CE-Kennzeichnung: vorhanden, keine Angabe zum Zeitpunkt; derzeit geltende Risikoklasse nach MDD laut Hersteller: 1)	55
Tabelle 3.3-10:	Aktuell laufende Studien zu M-sense (CE-Kennzeichnung: vorhanden, keine Angabe zum Zeitpunkt; derzeit geltende Risikoklasse nach MDD laut Hersteller: 1).....	56
Tabelle 3.3-11:	Publizierte Studien zu mySugr (CE-Kennzeichnung: vorhanden (2013?) <i>derzeit geltende Risikoklasse nach MDD laut Hersteller: 1, Bolusrechner als Modul Klasse 2b)</i>	57
Tabelle 3.3-12:	Digitale Gesundheitsanwendung, Risikoklasse und höchste Evidenzstufe	59
Tabelle 3.4-1:	Risikoklassen nach 2017/745 MDR und Evidenzstufen nach NICE.....	61
Tabelle 3.4-2:	Einteilung der digitalen Gesundheitsanwendung nach derzeitiger Risikoklasse, Risikoklasse nach 2017/745 und Evidenzstufen nach NICE.....	62
Tabelle 3.4-3:	Potentielle Empfehlung ausgewählter digitaler Gesundheitsanwendungen unter Anwendung von [28] basierend auf vorliegender Evidenz und relevante Bewertungsaspekte	64
Tabelle 3.5-1:	In Österreich angebotene Apps (2018)	78
Table 7-1:	A module which can be used in addition to existing core HTA domains to adapt HTA and/or reimbursement evaluation frameworks to assess MMAs	99
Table 7-2:	HTA domains and criteria of the Digi-HTA framework.....	104
Table 7-3:	Overview of the DEDHI framework	106
Table 7-4:	Aussagen internationaler HTA-Institutionen zu Assessment Frameworks für digitale Gesundheitsanwendungen	109

Abkürzungsverzeichnis

A.I.....	Artificial Intelligence
AHTA.....	Adelaide Health Technology Assessment
AIHTA	Austrian Institute for Health Technology Assessment
App.....	Application
AquAS	Agency of Health Quality and Assessment of Catalonia
AT.....	Österreich
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte (Deutschland)
BMI	Body-Mass-Index
CADTH	Canadian Agency for Drugs and Technologies in Health
CCG	Clinical Commissioning Groups (England)
CE.....	Conformité Européenne („Europäische Konformität“)
CED	Coverage with evidence development (Frankreich)
CHARISMHA	Chances and Risks of Mobile Health Apps
CHEATS.....	Generic Information Communication Technology Evaluation Framework
CNEDiMTS.....	Nationales Komitee zur Evaluierung von Medizinprodukten (Frankreich)
CONITEC	National Committee for Technology Incorporation (Brasilien)
CONSORT	Consolidated Standards of Reporting Trials
CONSORT-AI.....	Consolidated Standards of Reporting Trials-Artificial Intelligence
COPD.....	chronisch obstruktive Lungenerkrankung
CRD	Centre for Reviews and Dissemination
DEDHI	Design and Evaluation Framework for Digital Health Interventions
DES-SF.....	Diabetes Empowerment Scale-Short Form
DHI	Digital Health Intervention
DiGA.....	Digitale Gesundheitsanwendung
DiGAV	Digitale-Gesundheitsanwendungen-Verordnung
DiGa-Verzeichnis....	Verzeichnis für digitale Gesundheitsanwendungen
Digi-HTA	Digital Health Technology Assessment Framework
DM	Diabetes Mellitus
DNVF	Deutsches Netzwerk für Versorgungsforschung
DRG	Diagnosis Related Groups (diagnosebezogene Fallgruppen)
DVG	Digitale-Versorgung-Gesetz
eHealth.....	electronic Health
EKG.....	Elektrokardiogramm
EMA	ecological momentary assessment
ESIP	European Social Insurance Platform
ETAPES.....	Expérimentation de Télémédecine pour l'Amélioration des Parcours en Santé (Frankreich)
EU IVDR	European Union In-Vitro Diagnostics Regulation
EU MDR.....	European Union Medical Device Regulation
EUnetHTA	European Network for Health Technology Assessment
F&E.....	Forschung & Entwicklung
FAMHP	Federal Agency for Medicines and Health Products (Belgien)

FDA.....	Food and Drug Administration (USA)
FF.....	Forschungsfrage
FI-STAR.....	Future Internet Social and Technological Alignment Research
G-BA.....	Gemeinsamer Bundesausschuss (Deutschland)
GDPR.....	General Data Protection Regulation
GesÖk.....	Gesundheitsökonomie
GKV.....	Gesetzliche Krankenversicherung
HAD.....	Health Assessment Division of the Ministry of Public Health (Uruguay)
HAS.....	Haute Autorité de Santé (Frankreich)
HIH.....	Health Innovation Hub
HIS.....	Healthcare Improvement Scotland
HTA.....	Health Technology Assessment
HTW.....	Health Technology Wales
ICT.....	Information and Communication Technology
IDEAS.....	Integrate, Design, Assess and Share Framework
IfGP.....	Institut für Gesundheitsförderung und Prävention (Österreich)
IGES.....	Institut für Gesundheits- und Sozialforschung
IMDRF.....	International Medical Device Regulators Forum
INAHTA.....	International Network of Agencies for Health Technology Assessment
INAMI.....	Nationales Institut für Kranken- und Behindertenversicherung (Belgien)
INESSS.....	Institut National d'Excellence en Santé et en Services Sociaux (Kanada)
IQWiG.....	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Deutschland)
k.A.	keine Angabe
K.I.....	künstliche Intelligenz
KCE.....	Health Care Knowledge Centre (Belgien)
KVT.....	kognitive Verhaltenstherapie
LPPR.....	Liste des Produits et Prestations Remboursables (französische Erstattungsliste)
MARS.....	Mobile App Rating Scale
MAST.....	Model for Assessment of Telemedicine Applications
MAUQ.....	mHealth App Usability Questionnaire
MDCG.....	Medical Device Coordination Group
MDD.....	Medical Device Directive
MDSW.....	Medical Device Software
MeSH.....	Medical Subject Headings
mHealth.....	mobile Health
mind.....	Mindestens
MMA.....	Mobile Medical Application
mn.....	multinational
MOST.....	Multiphase Optimization Strategy Framework
MRT.....	micro-randomized trials
MTEP.....	Medical Technologies Evaluation Programme (England)
MTRC.....	MedTech Reimbursement Consulting
NeLL.....	National eHealth Living Lab (Niederlande)
NHS England.....	National Health Service England

NICE.....	National Institute for Health and Care Excellence
NICTIZ.....	Centre of Expertise for Standardization and eHealth (Niederlande)
NIHDI	Nationales Institut für Gesundheits- und Behindertenversicherung (Belgien)
NZa	Dutch Healthcare Authority
ORCHA	Organization for the Review of Care and Health Applications Framework
PHRC.....	Program for Hospital Clinical Research (Frankreich)
PICO-Schema.....	„Population, Intervention, Control, Outcome“-Schema
PRISMA.....	Transparent Reporting of Systematic Reviews and Meta-Analyses
PRME.....	Program for Medical Economic Research (Frankreich)
PROMIS-10	Patient-reported Outcomes Measurement Information System
PZ.....	Partena Ziekenfondsen (belgische Krankenkasse)
QALY	quality adjusted life years (qualitätsadjustierte Lebensjahre)
RCT.....	Randomized controlled trial (randomisiert kontrollierte Studie)
RIZIV	Nationales Institut für Kranken- und Invalidenversicherung (Belgien)
RK	Risikoklasse
RVS	Council of Public Health & Society (Niederlande)
RWE.....	real world evidence
SaMD	Software as Medical Device
SBU	Swedish Agency for Health Technology Assessment and Assessment of Social Services
SEC	Department of HTA at the State Expert Centre of the Ministry of Health (Ukraine)
SFOPH.....	Swiss Federal Office of Public Health
SMART	Sequential Multiple Assignment Randomized Trial
SMARTGEM.....	smartphonegestützte Migränetherapie
SOC	standard-of-care
SR.....	Systematic Review
technologiespez.	technologiespezifisch
TEMPEST	Technology, Economic, Market, Political, Evaluation, Social and Transformation Framework
TGA.....	Therapeutic Goods Administration (Australien)
TIDieR.....	Template for Intervention Description and Replication
TMNMT	tailor-made notched music training
TTP	time to pregnancy
UVON	Unified eValuation using Ontology Framework
UVT.....	Health Technology Assessment Unit (Italien)
WHO	World Health Organization

Zusammenfassung

Hintergrund

Der zunehmende Einsatz digitaler Technologien im Gesundheitswesen stellt Entscheidungsträger*innen vor neue Herausforderungen in der Evaluation dieser Applikationen. Für die Mehrzahl der verfügbaren digitalen Gesundheitsanwendungen liegt bislang keine oder wenig Evidenz zum (medizinischen, organisatorischen oder ökonomischem) Nutzen vor.

Bisher verfügbare Assessment Frameworks umfassen häufig nicht alle Domänen eines vollen HTAs. Gleichzeitig erfordert die gründliche Evaluation einer digitalen Gesundheitsanwendung zusätzliche Aspekte, die durch einen klassischen HTA-Bericht nicht abgedeckt sind. Dies sind technologiespezifische Aspekte wie beispielsweise Software-Updates, Fragen der Konnektivität und Kompatibilität sowie digitaler Datenschutz. Die vorliegende Arbeit zielt darauf ab, die Anwendbarkeit der bisher verfügbaren Assessment Frameworks auf ausgewählte digitale Gesundheitsanwendungen zu prüfen und Empfehlungen für Österreich zu entwickeln.

Evaluation digitaler Gesundheitsanwendungen herausfordernd

**zusätzliche zu HTA Domänen:
technologiespezifische Aspekte wie Updates, Konnektivität, Kompatibilität und Datenschutz**

Methoden

Zur Identifikation der Assessment Frameworks wurde eine systematische Literatursuche in vier Datenbanken und eine Befragung von INAHTA-Mitgliedern durchgeführt. Sechs Assessment Frameworks wurden exemplarisch ausgewählt und auf Ähnlichkeiten und Unterschiede analysiert. Zur Identifikation von bereits refundierten oder empfohlenen digitalen Gesundheitsanwendungen wurde eine Handsuche durchgeführt, ebenso wie zur Beschreibung der Strategien und Regelungen einzelner Länder im Umgang mit digitalen Gesundheitsanwendungen.

**Frameworks (FW):
systematische Suche in 4 Datenbanken & INAHTA-Befragung, Handsuche nach refundierten DiGA sowie nach Länder-Regelungen**

Anhand von Ein- und Ausschlusskriterien wurde eine exemplarische Auswahl von elf digitalen Gesundheitsanwendungen getroffen und die jeweiligen publizierten Studien durch eine Handsuche identifiziert. Die ausgewählten digitalen Gesundheitsanwendungen wurden in die voraussichtlichen Risikoklassen (nach 2017/745 EU MDR) eingeteilt. Danach wurde für die elf digitalen Gesundheitsanwendungen verglichen, ob die publizierten Studiendesigns den gewünschten Studiendesigns nach dem NICE Evidence Standards Framework entsprechen. Zuletzt wurden die relevanten Aspekte für eine HTA-Bewertung beschrieben.

Auswahl von 6 FW & 11 DiGA, Einteilung der DiGA in Risikoklassen, Abgleich der Studiendesigns mit Evidenzstandards im NICE FW

Ergebnisse

Vorschläge zu Studiendesigns wurden in 4 von 6 analysierten Assessment Frameworks genannt. Eine Risikoklassifikation der zu bewertenden digitalen Gesundheitsanwendung wurde in 1 von 6 analysierten Assessment Frameworks vorgeschlagen. Aspekte der künstlichen Intelligenz wurden durch 1 von 6 analysierten Frameworks berücksichtigt.

Unterschiede der FW in Bezug auf Studiendesigns, Risikoklassifikation und künstliche Intelligenz

Die höchste Evidenzstufe der jeweiligen digitalen Gesundheitsanwendungen reichte von (retrospektiven oder prospektiven) Datenbankanalysen von Nutzer- und Nutzungsdaten (2/11) zu Fall-Kontroll-Studien (3/11) bis randomisiert kontrollierte Studien (3/11). Auch wurden digitale Gesundheitsanwendungen ohne abgeschlossene Studien (3/11) gefunden.

Evidenzstufen: Datenbankanalysen, Fall-Kontroll-Studien, RCTs

**Unterschiede in
Risikoklassen und
gewählten Studiendesigns**

Die ausgewählten digitalen Gesundheitsanwendungen werden bereits von einzelnen Krankenkassen refundiert. Sie befinden sich nach Anwendung 2017/745 EU MDR in Risikoklasse mind. 2a (7/11) oder Risikoklasse mind. 2b (4/11). Von den ausgewählten digitalen Gesundheitsanwendungen entsprach nur für zwei Anwendungen das vorliegende Studiendesign den Evidenzstandards des NICE Frameworks (2/11), für eine Anwendung konnte keine eindeutige Aussage getroffen werden (1/11) und für acht Anwendungen entsprach das vorliegende Studiendesign nicht den Evidenzstandards des NICE Frameworks (8/11).

**Entwicklung zu
nationalen Regelungen**

Belgien, Frankreich, Niederlande, England und Deutschland weisen unterschiedliche Strategien zur Regelung und Refundierung digitaler Gesundheitsanwendungen auf. Rezente Entwicklungen in diesen Ländern zeigen, dass es allorts Bestrebungen gibt, Regelungen auf nationaler Ebene zu finden.

Schlussfolgerung

**Heterogenität der
Frameworks.
definierte Risikoklassen
nur durch NICE Framework,
laufende Studien machen
Hoffnung**

Es besteht eine große Heterogenität der Assessment Frameworks, vor allem in Bezug auf Einteilung in Risikoklassen sowie in Bezug auf technologiespezifische Aspekte wie Datenschutz und künstliche Intelligenz. Nur das Evidence Standards Framework von NICE bietet eine genaue Einteilung in definierte Risikoklassen mit entsprechend geforderten Studiendesigns. Die publizierte Evidenz der digitalen Gesundheitsanwendungen zeigt Unterschiede in der Wahl der Studiendesigns. Kontrollierte Studien mit Vergleich zum Versorgungsstandard sind nicht immer, aber in rezenten (laufenden) Studien häufiger gegeben.

Executive Summary

Background

The increasing use of digital technologies in the healthcare sector creates new challenges for decision-makers in the evaluation of these applications. For the majority of available digital health applications there is little or no evidence of (medical, organisational, economic) benefit.

Currently available assessment frameworks often do not cover all domains of a full HTA. At the same time, the thorough evaluation of a digital health application requires additional aspects that are not covered by a classic HTA report. These are technology-specific aspects such as software updates, connectivity and compatibility issues, as well as data protection and privacy. The aim of this report was to test the applicability of the available assessment frameworks on selected digital health applications and to develop recommendations for Austria.

challenges in evaluation of digital health applications

additional to HTA domains technology-specific aspects such as updates, connectivity, compatibility & data protection

Methods

To identify the assessment frameworks, a systematic literature search in four databases and a survey of INAHTA members was conducted. Six assessment frameworks were selected as examples and analyzed for similarities and differences. A hand search was conducted to identify already refunded or recommended digital health applications, and to describe the strategies and regulations of individual countries in dealing with digital health applications.

On the basis of inclusion and exclusion criteria, an exemplary selection of eleven digital health applications was made and the respective published studies were identified by a hand search. The selected digital health applications were classified into the probable risk classes (according to 2017/745 EU MDR). The eleven digital health applications were then compared to see whether the published study designs were in line with the desired study designs according to the NICE Evidence Standards Framework. Finally, the relevant aspects for an HTA evaluation were described.

frameworks (FW): systematic search in 4 databases, INAHTA survey, hand search for refunded mHealth Apps & for country regulations

selection of 6 FW & 11 Apps, classification of Apps into risk classes, comparison of study designs with NICE evidence standards

Results

Suggestions for study designs were mentioned in 4 out of 6 analysed assessment frameworks. A risk classification of the digital health application to be evaluated was suggested in 1 out of 6 analysed assessment frameworks. Aspects of artificial intelligence were examined in 1 out of 6 analyzed frameworks.

The highest evidence level of the respective digital health application ranged from (retrospective and prospective) database analyses of user data (2/11) to case-control studies (3/11) to randomized controlled trials (3/11). For some of the analyzed digital health applications there were no completed studies found (3/11).

The selected digital health applications are reimbursed by selective contracts of individual health insurance companies. After application of 2017/745 EU MDR, the selected digital health applications are in risk class of 2a or higher (7/11) or risk class 2b or higher (4/11). Of the selected digital health applications, only two applications have evidence that matches the required study

FW differ in terms of study design, risk classification and artificial intelligence

evidence: database analyses, case-control studies, RCTs

differences in risk classes and chosen study designs

design of the NICE evidence standards framework (2/11). For one application no clear statement could be made (1/11) and for eight applications the present study design did not meet the evidence standards of the NICE framework (8/11).

regulations on national level in development

Belgium, France, the Netherlands, England and Germany have different strategies for regulating and refunding digital health applications. Recent developments in these countries show that there are efforts everywhere to find regulations at national level.

Conclusion

heterogeneity of the frameworks
only NICE Framework defines risk classes, ongoing studies give hope for better evidence

There is a great heterogeneity of assessment frameworks, especially in terms of risk classification and technology-specific aspects such as data protection and artificial intelligence. Only the Evidence Standards Framework of NICE offers a precise classification into defined risk classes with correspondingly required study designs. The evidence of digital health applications shows differences in the choice of study designs. Controlled studies with comparison to standard of care are not always available, but are more common in recent (ongoing) studies.

1 Hintergrund

1.1 Einleitung

Digitale Technologien sind Teil unseres Alltags. Dabei hat in den letzten Jahren insbesondere die Nutzung mobiler digitaler Technologien an Bedeutung gewonnen und macht inzwischen den Großteil der Nutzung digitaler Technologien aus (vgl. Abbildung 1.1-1).

digitale Technologien sind Teil des Alltags

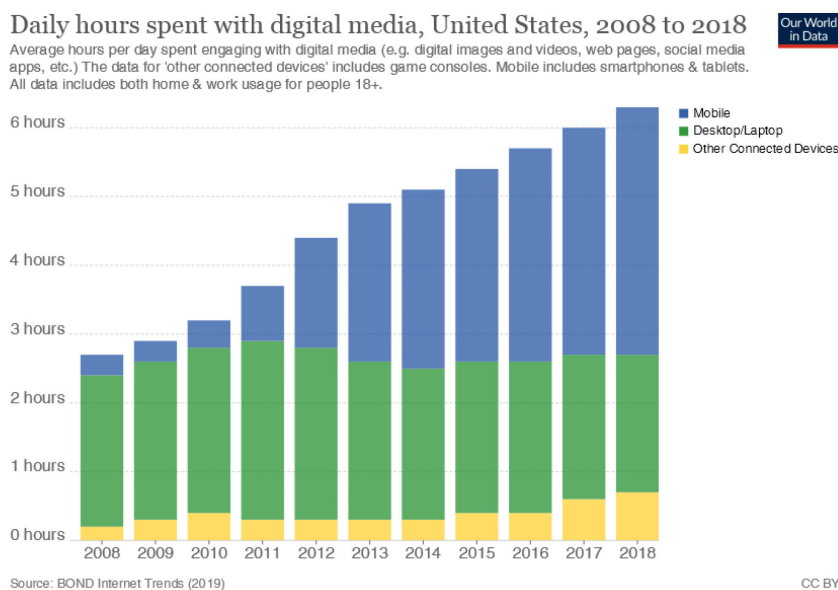


Abbildung 1.1-1: Durchschnittliche Tagesnutzung digitaler Technologien in den USA 2008-2018 [1]

Diese vermehrte Nutzung mobiler digitaler Technologien betrifft auch den Gesundheitsbereich. Eine im Mai 2019 von Bitkom Research in Deutschland durchgeführte Umfrage an 792 Teilnehmer*innen ergab, dass zu diesem Zeitpunkt etwa zwei von drei Smartphone-Besitzer*innen Gesundheits-Apps nutzen. Dazu zählen Apps für Sportübungen, Apps zum Aufzeichnen von Körperfunktionen wie Herzfrequenz oder Schritten, Apps mit Informationen zu Gesundheitsthemen, Erinnerungs-Apps für die Einnahme von Medikamenten oder zur Durchführung von Impfungen, sowie Apps zum Thema psychische Gesundheit und Stressreduktion [2]. Eine weitere im Juli 2020 von Bitkom Research durchgeführte Umfrage an 909 Teilnehmer*innen zeigte, dass die Nutzung weiter zugenommen hat und inzwischen etwa drei von vier Smartphone-Besitzer*innen Gesundheits-Apps nutzen. Dabei werden – unter Covid-19 – Video-Sprechstunden mit den Ärzt*innen und Therapeut*innen zunehmend beliebter, unter anderem aus Sorge, sich im Wartezimmer mit anderen Infektionen zu infizieren [3].

auch im Gesundheitsbereich

2 Befragungen aus Deutschland (2019, 2020):

**2019:
2 von 3 Smartphone Besitzern verwenden GesundheitsApps**

**2020:
bereits 3 von 4**

große Erwartungen, aber auch neue Herausforderungen in der Evaluation des tatsächlichen Nutzens	Das Angebot von digitalen Technologien im Gesundheitsbereich ist dabei riesig und reicht von rein informativen Anwendungen über Anwendungen mit diagnostischen Funktionen sowie Anwendungen für therapeutische Zwecke. Da mit diesen digitalen Gesundheitsanwendungen große Erwartungen verbunden sind, stellt dieser zunehmende Einsatz digitaler Technologien [4, 5] Entscheidungsträger*innen im Gesundheitswesen vor neue Herausforderungen in der Evaluation dieser Applikationen. Digitale Gesundheitsanwendungen haben das Potenzial, das Monitoring und das Management von Krankheiten zu verbessern. Durch telemedizinische Dienste und Telemonitoring könnten unnötige Wege für Patient*innen und Therapeut*innen vermieden werden. Durch eine Fernüberwachung, insbesondere bei chronischen Erkrankungen, könnten Krankenhausaufenthalte vermieden werden. Bei kontinuierlich gemessenen Vitalparametern könnte eine Entgleisung eines Werts rasch erkannt und entsprechende Maßnahmen in Gang gebracht werden.
Potenziale durch Früherkennung und Monitoring, Vermeidung von Krankenhausaufenthalten	
aber auch Risiken: ständiges Monitoren und Vorhersagen von Krankheiten, Kontrolle	Für die Mehrzahl der verfügbaren Gesundheits-Apps liegt bislang aber keine oder wenig Evidenz zum tatsächlichen Nutzen vor [6, 7]. Darüber hinaus können von digitalen Gesundheitsanwendungen auch diverse Risiken ausgehen. Eine ständige Überwachung von Gesundheitsdaten mit automatisierten Vorhersagen, Erkennung von Krankheiten sowie Belohnungs- und Strafsystemen sind kritisch zu beurteilen. So wird beispielsweise in Südkorea im Zuge der Covid-19-Pandemie eine Gesundheits-App eingesetzt, welche über den Standort der Nutzer*innen überprüft, ob sie sich an die Quarantäne halten [8]. In der Befragung von Bitkom Research 2020 an 682 Nutzer*innen von Gesundheits-Apps gaben 18 % an, sich durch die Nutzung dieser Apps häufig unter Druck gesetzt zu fühlen [3].
Nutzer*innen fühlen sich unter Druck	
Fehlinformationen: Fehldiagnosen und –behandlungen als Folge	Weitere Risiken von digitalen Gesundheitsanwendungen sind mögliche Nebenwirkungen oder das Ausbleiben der erwünschten Wirkung. Da die digitalen Gesundheitsanwendungen Informationen vermitteln, bestehen Gefahren bei Fehlinformationen. Fehldiagnosen oder Fehlbehandlungen sind mögliche Konsequenzen. Weitere Probleme können durch eine fehlerhafte Bedienung der digitalen Gesundheitsanwendung durch die Nutzer*innen entstehen, etwa bei Design-Problemen der Anwendungen, oder wenn diese außerhalb ihres bestimmten Einsatzbereiches verwendet werden [9].
Anwendungsfehler	
IGES 2016: Empfehlung zur stärkeren Regulierung & infolge Bezahlung von geeigneten Apps DE: DVG	Ein im September 2016 veröffentlichter Bericht des IGES-Instituts sprach sich für eine stärkere Regulierung von medizinischer Software aus. Darauf aufbauend könne die Implementierung einer generellen, nationalen Refundierung von geeigneten digitalen Gesundheitsanwendung erfolgen. Diese Refundierungsmodelle können ähnlich zu den Modellen für Pharmazeutika oder Medizinprodukte gestaltet sein, d. h. auf Nutzen und Sicherheit überprüfen [10]. In Deutschland soll das Digitale-Versorgung-Gesetz (DVG) Ärzt*innen und Therapeut*innen künftig eine solche Verschreibung und die Refundierung von Gesundheits-Apps ermöglichen [11]. Bereits jetzt werden erste Gesundheits-Apps als medizinische Leistung von einzelnen (deutschen, belgischen, niederländischen, aber auch österreichischen) Krankenkassen übernommen.
SV & KV in DE, BE, NL, Ö übernehmen bereits Kosten für Apps	
verfügbare Frameworks: limitierte Evidenz zur Anwendbarkeit	Bisher verfügbare Frameworks zum Assessment von digitalen Gesundheitsanwendungen umfassen häufig nicht alle Domänen eines vollen HTAs. Gleichzeitig erfordert die gründliche Evaluation einer digitalen Gesundheitsanwendung aber zusätzliche Domänen, die durch einen klassischen HTA-Bericht nicht abgedeckt sind (technologie-spezifische Aspekte wie beispielsweise Software-Updates, Fragen der Konnektivität und Kompatibilität sowie digitaler Datenschutz) [12]. Ein systematisches Review zu HTA Frameworks
HTA-Framework als Basis, aber nicht ausreichend	

für eHealth Anwendungen identifizierte und analysierte 21 Assessment Frameworks. Dabei geben die Autor*innen an dass die Evidenz der Anwendbarkeit dieser Frameworks limitiert ist [13].

Der Bedarf eines standardisierten Bewertungsinstruments zur transparenten Evaluation digitaler Gesundheitsanwendungen ist hoch. Nicht nur Entscheidungsträger*innen im Gesundheitswesen, sondern auch Entwickler*innen dieser Technologien und nicht zuletzt Patient*innen können ein solches Bewertungsinstrument nutzen und sich daran orientieren, welcher Nutzenbeleg für eine öffentliche Kostenerstattung vonnöten ist. Das Ziel dieser vorliegenden Arbeit ist eine Orientierungshilfe für Refundierungsinstitutionen, aber auch für Hersteller*innen, welche Evidenz für welche Art von digitaler Gesundheitsanwendung erwartet wird. Auch wenn sich digitale Gesundheitsanwendungen von konventionellen Interventionen im Gesundheitswesen unterscheiden, sollten sie dabei dieselben Evidenzanforderungen zum Nutzen, möglichen Schäden und entstehenden Kosten erfüllen.

Ziel des vorliegenden Berichts:

Orientierungshilfe für Zahler, Hersteller, aber auch Patient*innen

1.2 Definitionen

Es existiert eine Vielzahl an Bezeichnungen, Grundbegriffen und Termini für Technologien im Gesundheitsbereich, wobei exakte Definitionen bislang nicht vorhanden sind [14]. Um für ein standardisiertes Vokabular in der Beschreibung dieser Technologien zu sorgen, schlug die WHO 2018 eine Taxonomie vor, um verschiedene digitale und mobile Technologien gezielter zu kategorisieren. Dabei unterscheidet die *Classification of Digital Health Interventions* Taxonomie grob vier Bereiche, je nachdem ob die Technologie von den Anwender*innen bzw. Patient*innen oder durch die Ärzt*innen bzw. Therapeut*innen verwendet wird. Technologien für Datendienste sowie Technologien die dem Management im Gesundheitswesen auf Systemebene dienen, bilden jeweils eigene Kategorien (vgl. Abbildung 1.2-1).

Vielzahl an Bezeichnungen ohne exakte Definitionen

WHO 2018: Taxonomie zur Kategorisierung

Die meisten Anwendungen mit den häufig verwendeten Bezeichnungen „Gesundheits-Apps“ oder „digitale Gesundheitsanwendungen“ sind demnach der Kategorie „1.0 Clients“ zuzuordnen, da sie für die Gesundheit der Anwender*innen eingesetzt und von den Anwender*innen selbst bedient werden. Demgegenüber stehen Anwendungen auf Gesundheitsanbieter-Ebene oder Anwendungen auf der Gesundheitssystem-Ebene.

digitale Gesundheitsanwendungen von Anwender*innen selbst bedient

Definitionen:

- *eHealth (electronic Health)*: ist definiert als „der Einsatz von Informations- und Kommunikationstechnologie zur Unterstützung der Gesundheit und gesundheitsbezogener Bereiche“ [16].
- *mHealth (mobile Health)*: ist eine Komponente von eHealth und definiert als „die Nutzung mobiler drahtloser Technologie für die Gesundheit“ [16].
- *Digital Health*: ein weit gefasster Überbegriff, der eHealth (einschließlich mHealth) und neu entstehende Bereiche umfasst, wie die Anwendung der Informatik in den Bereichen künstliche Intelligenz, Big Data und Genomik [16].

Definitionen: eHealth, mHealth, Digital Health

Definitionen:
Gesundheits-Apps,
Medizin-Apps

Auf der Ebene einzelner Anwendungen werden häufig die Begriffe Gesundheits-App und Medizin-App verwendet, welche durch die AG Digital Health des deutschen Netzwerks für Versorgungsforschung (DNVF) wie folgt definiert werden [17]:

- **Gesundheits-Apps:** Anwendungen, die sich an Verbraucher richten. Sie sollen in präventiver Absicht genutzt werden oder helfen, den Lebensstil gesundheitsförderlich zu gestalten oder in diese Richtung zu verändern.
- **Medizin-Apps:** Anwendungen, die zur Unterstützung der „Selbstbefähigung“ (Empowerment) genutzt werden und bei der Bewältigung z. B. von chronischen Krankheiten oder in der Rehabilitation unterstützend eingesetzt werden [17].

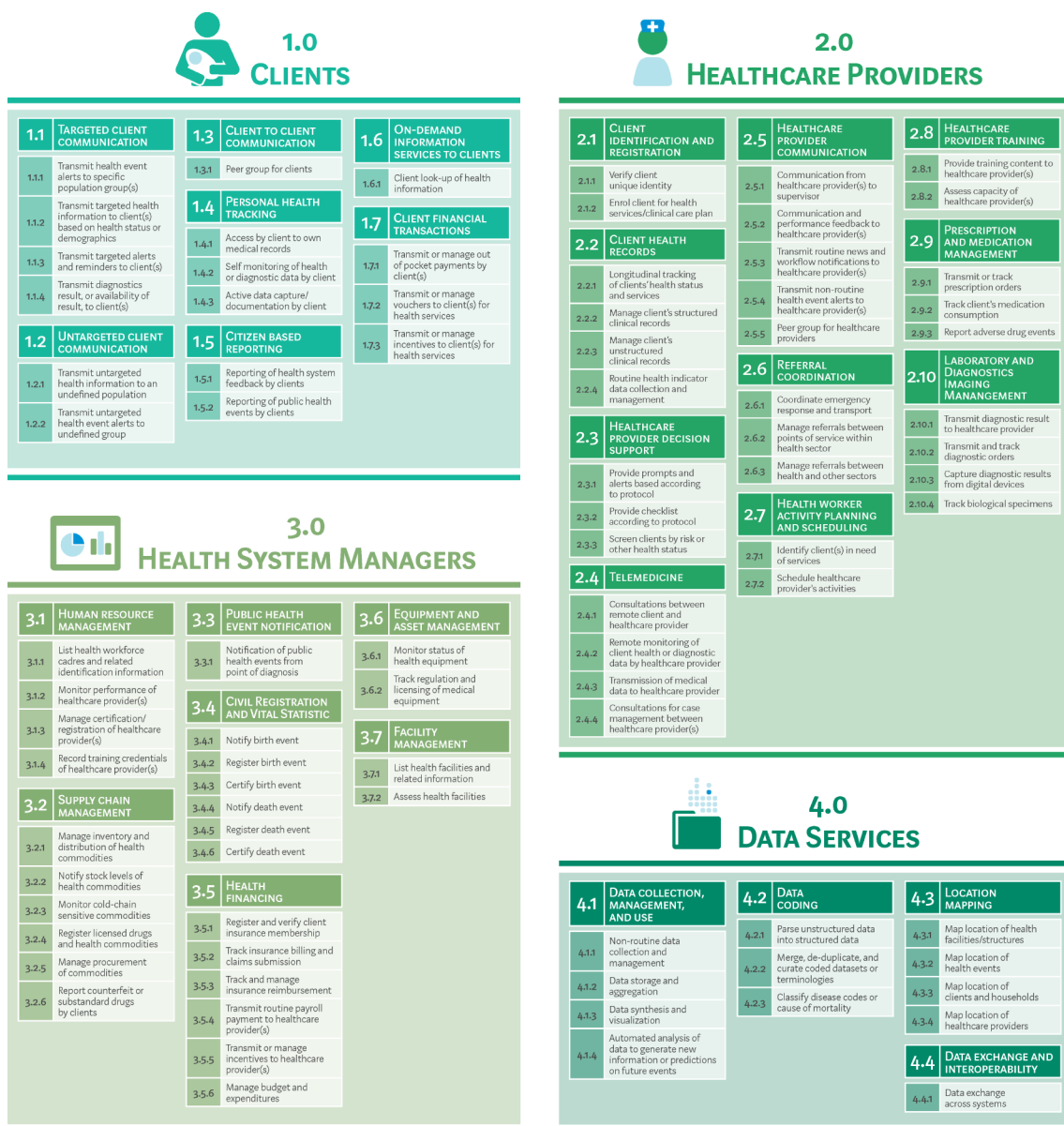


Abbildung 1.2-1: WHO Classification of digital health interventions v1.0 [15, 16]

1.3 Zulassung und Risikoklassifizierung von digitalen Gesundheitsanwendungen

2017 wurden die beiden neuen Verordnungen für Medizinprodukte (MDR) und In-vitro-Diagnostika (IVDR) durch das Europäische Parlament verabschiedet [18, 19]. Die verpflichtende Umsetzung der Verordnung sollte innerhalb von drei Jahren erfolgen, wurde aber aufgrund von Covid-19 auf Mai 2021 für Medizinprodukte und Mai 2022 für In-vitro-Diagnostika verschoben. Im Zuge der Implementierung der rechtlichen Rahmenbedingungen wurden nicht nur die Zulassungsbedingungen für unterschiedliche Risikoklassen von Medizinprodukten und In-vitro Diagnostika festgeschrieben, sondern auch (erste) Anleitungen („Guidance“) zu Definitionen von Medizinprodukten, die mit Software arbeiten [20] sowie deren klinische Evaluierung [21] festgelegt.

Die Europäischen Regulatoren arbeiten hier in engem Verbund mit internationalen Regulatoren (IMDRF, International Medical Device Regulators Forum). Die Medical Device Coordination Group (MDCG) definiert Medical Device Software (MDSW, Medizinproduktesoftware), die FDA als „Software as a Medical Device“ (SaMD) folgendermaßen

1. Software, die allein („standalone software“) oder
2. Software, in Kombination mit (Hardware-) Medizinprodukt oder In-Vitro Diagnostikum

einen beabsichtigten medizinischen Zweck verfolgt und auch andere Medizinprodukte steuert oder beeinflusst. Die Definition beinhaltet MDSW auch ungeachtet dessen, wo sie eingesetzt wird (Cloud, Computer, Mobiltelefon) und wer sie einsetzt (Gesundheitspersonal oder Laiennutzer*innen).

Beispiele für „standalone“ software sind:

- MDSW, die mütterliche Parameter wie Alter, Konzentration von Serummarkern und weitere Informationen, die durch fetale Ultraschalluntersuchung zur Bewertung des Trisomierisikos 21 gewonnen wurden, verwendet.
- MDSW, die Messungen von transrektalen Ultraschallbefunden, Alter und Ergebnissen von In-vitro-Diagnostik nutzt, um das Risiko eines Patienten, an Prostatakrebs zu erkranken, berechnet.
- MDSW Smartwatch-App, die Alarmmeldungen über den Gesundheitszustand an den Benutzer senden soll, wenn unregelmäßige Herzschläge im Zuge eines Monitoring von Herzrhythmusstörungen erkannt werden.

Beispiele für MDSW Kombinationsprodukte sind:

- MDSW zur Bildanalyse von Melanomen, die einen Nahinfrarot-Laserlichtscanner steuern soll.
- MDSW zur Messung und Übertragung des Blutzuckerspiegels, zur Berechnung der erforderlichen Insulindosis und zur Steuerung der Insulinpumpe zur Verabreichung der berechneten Dosis (closed loop systems).

Im Oktober 2019 veröffentlichte die MDCG eine Guidance (MDCG 2019-11: Guidance on Qualification and Classification of Software in Regulation [20]) basierend auf einem IMDRF Vorschlag 2017 [22] – eine eigene Klassifikation für die EU, die an die EU-Risikoklassen angepasst ist. Diese Risiko-Klassifizierung ist die Grundlage für spezifische Anforderungen an klinische Evi-

**2017:
neuen Verordnungen
für Medizinprodukte (MDR)
und In-vitro-Diagnostika
(IVDR)**

**verbindliche
Implementierung bis
Mai 2021 (MDR)/
Mai 2022 (IVDR)**

**EU-Regulatoren arbeiten
in enger Kooperation mit
internationalen
Regulatoren**

**EU: Medical Device
Software (MDSW)**

**Beispiele für
„standalone“ software**

**Beispiele für MDSW
Kombinationsprodukte**

**MDCG Guidance:
2019-11**

denz und entspricht der Signifikanz der Information für (klinische, gesundheitssystemische) Entscheidungen und den Impact auf Patient*innen. In drei Risikoklassen und neun Abstufungen werden die MDSW danach eingeteilt, ob sie

3 Risikoklassen und 9 Abstufungen zur Klassifizierung

- „nur“ ein klinisches Management informieren (in nicht-schwerwiegenden, schwerwiegenden oder kritischen Situationen oder medizinischen Zuständen) oder
- Entscheidungen bereits steuern (in nicht-schwerwiegenden, schwerwiegenden oder kritischen Situationen oder medizinischen Zuständen) oder
- gar selbständig behandeln oder diagnostizieren (in nicht-schwerwiegenden, schwerwiegenden oder kritischen Situationen oder medizinischen Zuständen).

Tabelle 1.3-1: Risikoklassen nach Signifikanz der Information für Entscheidungen, MDCG–Guidance [20]

		Significance of Information provided by the MDSW to a healthcare situation related to diagnosis/therapy		
State of Healthcare situation or patient condition		High Treat or diagnose ~ IMDRF 5.1.1	Medium Drives clinical management ~ IMDRF 5.1.2	Low Informs clinical management (everything else)
	Critical situation or patient condition ~ IMDRF 5.2.1	Class III Category IV.i	Class IIb Category III.i	Class IIa Category I.i
	Serious situation or patient condition ~ IMDRF 5.2.2	Class IIb Category III.ii	Class IIa Category II.ii	Class IIa Category I.ii
	Non-serious situation or patient condition (everything else)	Class IIa Category II.iii	Class IIa Category I.iii	Class IIa Category I.i

Regel 11 in diesem MDCG Dokument (MDCG–Guidance 2019-11 [20], S 13) besagt, dass

Regel 11 gibt Anleitung, wann

Risikoklasse IIa, IIb oder III

„Software intended to provide information which is used to take decisions with diagnosis or therapeutic purposes is classified as class IIa, except if such decisions have an impact that may cause:

- *death or an irreversible deterioration of a person’s state of health, in which case it is in class III; or*
- *a serious deterioration of a person’s state of health or a surgical intervention, in which case it is classified as class IIb.*

Software intended to monitor physiological processes is classified as class IIa, except if it is intended for monitoring of vital physiological parameters, where the nature of variations of those parameters is such that it could result in immediate danger to the patient, in which case it is classified as class IIb. All other software is classified as class I.”

Nach einer Einstufung der Risikoklasse ist die klinische Evidenz für MDSW durch drei Komponenten definiert (MDCG Guidance on Clinical Evaluation (MDR)/Performance Evaluation (IVDR) of Medical Device Software, 2020-1, März 2020, [21]):

- *Klinisch-wissenschaftliche Validität* – definiert durch den Nachweis oder eine gute Begründung, dass der Output des MDSW basierend auf den Inputs mit dem entsprechenden physiologischen Zustand oder der klinischen Indikation assoziiert ist.
- *Technische/Analytische Performanz* – definiert durch den Nachweis der Fähigkeit der MDSW, aus den Eingabedaten genaue und verlässliche Aussagen zu erzeugen.
- *Klinische Performanz/klinische Evaluation* – definiert durch den Nachweis der Fähigkeit der MDSW, klinisch relevante Aussagen in Übereinstimmung mit dem beabsichtigten Zweck zu machen. Die klinische Relevanz wird durch positive Auswirkungen auf die – mit der MDSW zusammenhängenden – Funktion bestimmt: verbesserte Gesundheit aufgrund messbarer, patientenrelevanter klinischer Ergebnisse in Vor-sorge, Überwachung, Diagnose oder Diagnoseunterstützung von Pati-ent*innen.

Die klinische Evaluation der MSDW kann sich auf verschiedene Datenquellen berufen: Daten aus Studien mit gleichem/äquivalentem MDSW, prospektive oder retrospektive Studien, Datenbanken- und Register-Studien, aber auch Routine- und Real-World-Performance Daten [21].

Zusammenfassend ist festzuhalten, dass die Europäischen Regulatoren eng mit anderen internationalen Regulatoren (FDA) bei der Definition von Medizinprodukten und deren Risikoklassen zusammenarbeiten. Bei den konkreten Evidenzanforderungen je Risikoklasse bleiben die Guidance-Papiere aber unspezifisch und vage, führen aber anhand von Beispielen unterschiedlicher MDSWs mögliche Datenquellen zur Bewertung der klinischen Performanz und Evaluation an [21]. Wie die MDR/IVDR in der Praxis gelebt werden wird, ist derzeit noch nicht abzuschätzen. Die vorgelegten Grundlagen eignen sich aber dafür, entsprechende Anforderungen an Evidenz für – nach Risikoklassen abgestufte – Nutzenbewertungen und Health Technology Assessments (HTAs) abzuleiten (vgl. dazu auch das NICE Framework, Kapitel 3.1.2).

**MDCG Guidance: 2020-1
3 Komponenten bei der
Evaluation**

**klinisch-wissenschaftliche
Validität**

**technische-analytische
Performanz**

**klinische Performanz
und Evaluation**

**Quellen für klinische
Evaluation:
Primärstudien, Register**

**letztendlich:
MDCG bleibt unspezifisch
zu konkreten
Evidenzanforderungen
für MDSW**

2 Methodisches Vorgehen

2.1 Projektziel

Es ist die Intention dieses Berichts, eine Orientierungshilfe für Refundierungsinstitutionen, aber auch für Hersteller*innen zu geben, welche Evidenz für welche Art von digitaler Gesundheitsanwendung erwartet wird. Dafür wird die Anwendbarkeit der bisher verfügbaren Frameworks auf digitale Gesundheitsanwendungen geprüft. Folgende Forschungsfragen sollen beantwortet werden:

**Orientierungshilfe
für Zahler:
Prüfung bisheriger
Frameworks**

2.2 Forschungsfragen

1. Welche internationalen Bewertungsinstrumente für digitale Gesundheitsanwendungen gibt es, die für Refundierungsentscheidungen verwendet werden können?
2. Welche digitalen Gesundheitsanwendungen werden von Sozial-/Krankenversicherungen in ausgewählten Ländern bereits bezahlt, resp. empfohlen?
3. Welche Evidenz zum Nutzen liegt für die bereits refundierten digitalen Gesundheitsanwendungen vor? Welche sollte entsprechend den internationalen Bewertungsinstrumenten vorliegen?
4. Wie ist die Anwendbarkeit ausgewählter Bewertungsinstrumente für digitale Gesundheitsanwendungen?
5. Welche Herangehensweisen, Strategien und Modelle zur Refundierung digitaler Gesundheitsanwendungen kommen in ausgewählten Ländern zum Einsatz?

**FF:
Welche Frameworks gibt es
und eignen sich diese
Frameworks zur Auswahl
für HTA-Bewertungen?**

**Welche digitalen Apps
werden schon bezahlt und
welche Evidenz gibt es für
diese Apps?**

**Wie gehen andere Länder
vor?**

2.3 Systematische Literatursuche

Für die Beantwortung der ersten Forschungsfrage (Bewertungsinstrumente für digitale Gesundheitsanwendungen) wurde im Juni 2020 eine systematische Literatursuche durchgeführt. Diese Literatursuche orientiert sich an einer von Vis et al. [13] durchgeführten systematischen Übersichtsarbeit zu Assessment Frameworks für digitale Gesundheitsanwendungen. Die systematische Literatursuche dient einer Aktualisierung des Reviews von Vis et al.. Die Literatursuche wurde in den folgenden vier Datenbanken durchgeführt:

**Methode FF1:
systematische Suche 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 0).

2.3.1 Ein- und Ausschlusskriterien

Einschlusskriterien	Folgende Einschlusskriterien wurden für die systematische Literatursuche gewählt: ■ Die Studie beschäftigt sich mit Bewertungsinstrumenten für digitale Gesundheitsanwendungen ■ Das beschriebene Bewertungsinstrument für digitale Gesundheitsanwendungen deckt mehrere (oder alle) Domänen eines HTA ab und umfasst weitere technologie-spezifische Aspekte ■ Die Studie ist in englischer oder deutscher Sprache publiziert ■ Publikationszeitraum: ab 01.03.2018
Ausschlusskriterien	Folgende Ausschlusskriterien wurden für die systematische Literatursuche gewählt: ■ Das primäre Thema der Studie ist nicht die Bewertung digitaler Gesundheitsanwendungen ■ Die Studie beschreibt Methoden zum Design bzw. zur Implementierung von digitalen Gesundheitsanwendungen ■ Die Studie verfügt über keine eindeutige Beschreibung des Kontexts ■ Die Studie ist nicht in englischer oder deutscher Sprache publiziert ■ Publikationszeitraum vor 28.02.2018
Einschränkung des Suchzeitraums: ab 01.03.2018	Zur Erklärung der Einschränkung des Suchzeitraums ab dem 01.03.2018: Einerseits wurde der Ansatz verfolgt sich in der systematischen Suche auf aktuelle Assessment Frameworks zu beschränken. Andererseits wurde bereits von Vis et al. [13] eine systematische Übersichtsarbeit zu Assessment Frameworks für digitale Gesundheitsanwendungen durchgeführt; in dieser Übersichtsarbeit wird die Literatursuche mit März 2018 angegeben; dieselbe Suchstrategie kam für diesen Bericht zur Anwendung.

2.3.2 Literaturauswahl

Literaturauswahl aus 814 Quellen	Insgesamt standen 814 Quellen für die Literaturauswahl zur Verfügung. Die Literatur wurde von zwei Personen (RJ, CW) unabhängig voneinander begutachtet. Differenzen wurden durch Diskussion und Konsens oder die Einbindung einer dritten Person gelöst. Für das Abstract Screening wurde die webbasierte Plattform Rayyan genutzt [23]. Der Auswahlprozess ist in Abbildung 2.3-1 dargestellt:
---	---

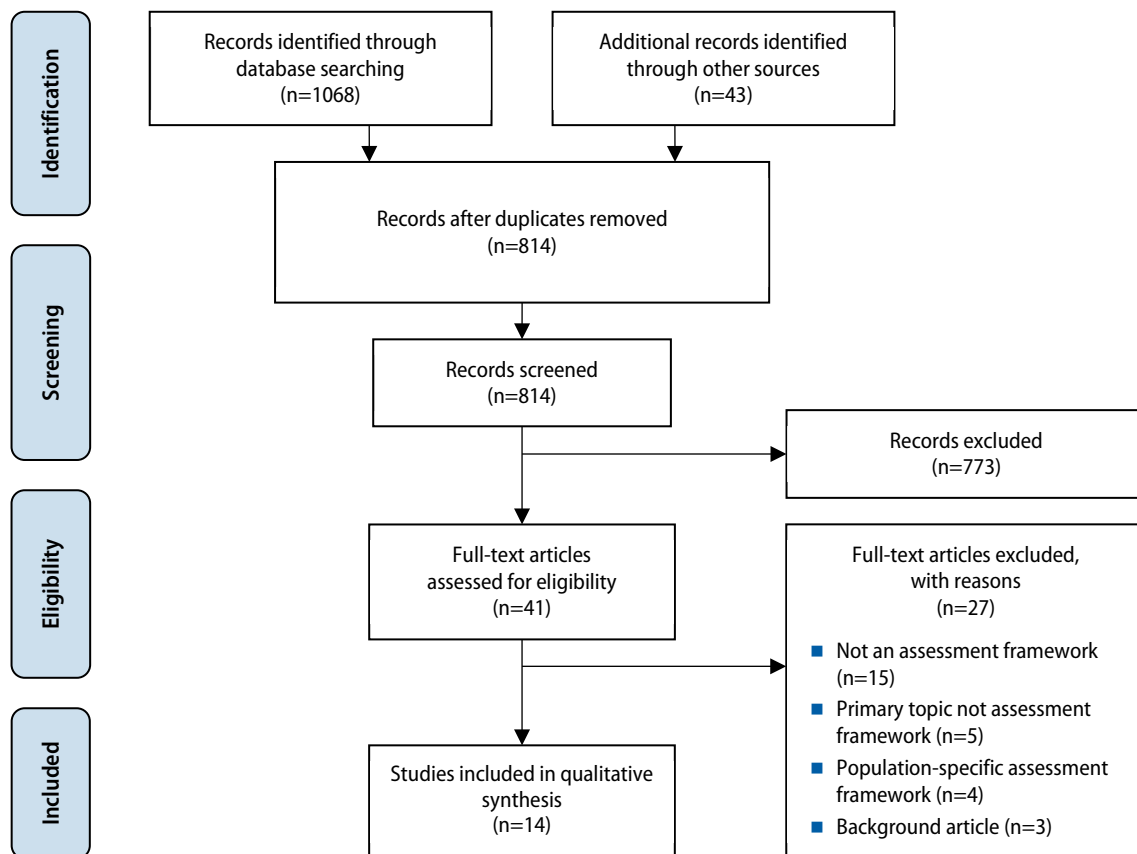


Abbildung 2.3-1: Darstellung des Auswahlprozesses (PRISMA Flow Diagram)

2.4 INAHTA ListServ zu Assessment Frameworks

Zur Ergänzung der systematischen Literatursuche wurde am 08.06.2020 eine Anfrage an INAHTA ListServ gestellt um weitere Assessment Frameworks für digitale Gesundheitsanwendungen zu identifizieren, und zu erfahren welche Assessment Frameworks bereits in der Anwendung genutzt werden. Folgende Fragestellung wurde gewählt:

The Austrian Institute for Health Technology Assessment (AIHTA) is in the process of reviewing frameworks for the assessment of digital health technologies. The intent is to evaluate the evidence on effectiveness and safety of reimbursed digital health technologies (in particular mobile health applications). To this aim, AIHTA would like to kindly ask INAHTA members for your brief response to the following questions:

- *What framework(s) does your organization consider suitable for use in HTA to assess digital health applications?*
- *Is your organization currently developing or adapting a framework to be used for assessing mobile health applications?*

Eine Zusammenfassung der INAHTA ListServ Antworten findet sich im Ergebniskapitel „Übersicht Assessment Frameworks“ (Kapitel 3.1). Die gesammelten Antworten der INAHTA Mitglieder finden sich im Appendix (Kapitel 7.2.1).

INAHTA ListServ Anfrage für weitere Assessment Frameworks

INAHTA Anfrage im Detail

2.5 Auswahl und Datenextraktion der Assessment Frameworks

Auswahl nach Aktualität, Ausgereiftheit und Anwendbarkeit

Ziel der Auswahl der Assessment Frameworks war eine exemplarische Darstellung verschiedener internationaler Bewertungsinstrumente für digitale Gesundheitsanwendungen. Kriterien bei der Auswahl der Assessment Frameworks waren Aktualität, Ausgereiftheit und zu erwartende Anwendbarkeit auf digitale Gesundheitsanwendungen.

etablierte eHealth Frameworks ohne Berücksichtigung von mHealth ausgeschlossen

Bereits etablierte und erprobte Frameworks wurden nicht gewählt, wenn diese zwar für den Bereich eHealth aufgestellt sind, jedoch die Entwicklungen im Bereich mHealth nicht mehr ausreichend berücksichtigen. Jene Frameworks die über die systematische Literatursuche und Handsuche identifiziert wurden, für die detaillierte Analyse jedoch nicht herangezogen wurden, sind im Kapitel 3.1.7 beschrieben.

Datenextraktion aus ausgewählten Assessment Frameworks

Bei der Auswahl der Assessment Frameworks wurden auch die von den IN-AHTA Mitgliedern empfohlenen Assessment Frameworks, die in Vis et al. [13] identifizierten sowie die nach 2018 publizierten Frameworks berücksichtigt. Aus den ausgewählten Assessment Frameworks wurden Daten zu folgenden Kriterien extrahiert:

- Organisation/Institution, Land
- Finanzierung
- Zeitpunkt der Veröffentlichung
- Zielgruppe
- Entwicklungsstatus
- Anwendungsbereich
- Berücksichtigung von Risikoklassen
- Evidenzerfordernisse
- Zeitpunkt der Anwendung/Intention des Frameworks
- Format des Frameworks
- Bewertungsdomänen
- Technologiespezifische Aspekte
- Involvierung Patient*innenvertretung
- Framework-Spezifika

2.6 Identifikation von digitalen Gesundheitsanwendungen

2.6.1 Handsuche

Methode FF2: Handsuche nach digitalen Gesundheitsanwendungen

Zur Beantwortung der zweiten Forschungsfrage wurde eine Handsuche durchgeführt. Zur Identifikation digitaler Gesundheitsanwendungen, die bereits refundiert werden, oder deren Refundierung beantragt wurde, wurden folgende Quellen (Reports und Websites) verwendet:

- Digital Healthcare Products Report des IGES Instituts [10]
- Reimbursement Landscape for Health Apps Report der MTRC [24]
- <https://mtrconsult.com/news>

- <https://mhealthbelgium.be/>
- <https://hih-2025.de/>
- <https://www.nhs.uk/apps-library>
- Websites folgender deutscher Krankenversicherungen:
Techniker Krankenkasse, AOK Gesundheitskasse, IKK Classic,
Generali Krankenversicherung, Barmer Ersatzkasse, HUK-COBURG
Krankenversicherung, Postbeamtenkrankenkasse, DAK-Gesundheit

Zur Beantwortung der dritten Forschungsfrage wurde eine Handsuche durchgeführt. Zur Identifikation der CE-Kennzeichnung, der angegebenen Medizinproduktklasse und der verfügbaren Studien zu den ausgewählten digitalen Gesundheitsanwendungen wurden die Websites der digitalen Gesundheitsanwendungen für die Suche herangezogen. Ergänzend wurde nach aktuell laufenden Studien in der Studiendatenbank <https://clinicaltrials.gov/> gesucht.

Methode FF3:
Handsuche nach
CE-Kennzeichnung,
Risikoklasse und
verfügbaren Studien

2.6.2 Ein- und Ausschlusskriterien der digitalen Gesundheitsanwendungen

Folgende Einschlusskriterien wurden bei der Auswahl der digitalen Gesundheitsanwendungen gewählt:

- Die digitale Gesundheitsanwendung verfügt über eine CE-Kennzeichnung.
- Die digitale Gesundheitsanwendung wird durch einzelne Krankenkassen in ausgewählten Ländern bereits refundiert, eine Refundierung der digitalen Gesundheitsanwendung ist beantragt oder wird empfohlen.
- Die digitale Gesundheitsanwendung ist nach der EU-Verordnung MDR 2017/745 in einer höheren Medizinproduktrisikoklasse (mindestens 2a) einzuteilen. Digitale Gesundheitsanwendungen, die sich nach derzeitig geltender Risikoklassifizierung in Risikoklasse 1 befinden, bei Anwendung der MDR 2017/745 jedoch in der Risikoklassifizierung höhergestuft werden, sind ebenfalls eingeschlossen.
- Die digitale Gesundheitsanwendung ist in englischer oder deutscher Sprache verfügbar.
- Die digitale Gesundheitsanwendung ist im österreichischen Apple App Store und/oder Google Play Store verfügbar.

Einschlusskriterien:

CE-Kennzeichnung

Refundierung oder
Empfehlung
nach MDR 2017

mind. in
Risikoklasse 2a

englisch/deutsch

Folgende Ausschlusskriterien wurden bei der Auswahl der digitalen Gesundheitsanwendungen gewählt:

- Die digitale Gesundheitsanwendung verfügt über keine CE-Kennzeichnung.
- Die digitale Gesundheitsanwendung ist eher im Bereich Fitness-App als im Bereich Medizin-App einzuordnen.
- Die digitale Gesundheitsanwendung ist in die Risikoklasse 1 einzuteilen (nach Verordnung MDR 2017/745).
- Die digitale Gesundheitsanwendung ist nicht in englischer oder deutscher Sprache verfügbar.
- Die digitale Gesundheitsanwendung ist nicht im österreichischen Apple App Store oder Google Play Store verfügbar.

Ausschlusskriterien:

keine CE-Kennzeichnung
Risikoklasse 1

2.7 Auswahl der digitalen Gesundheitsanwendungen

Auswahl nach Funktionen, Risikoklasse, medizinischem Fachgebiet und Refundierungsmodell

Von den durch die Handsuche identifizierten digitalen Gesundheitsanwendungen wurden beispielhaft elf digitale Gesundheitsanwendungen für eine detailliertere Analyse ausgewählt. Bei der Auswahl der digitalen Gesundheitsanwendungen wurde darauf geachtet, dass diese verschiedene diagnostische und/oder therapeutische Funktionen anbieten, einer höheren Risikoklasse zuzuordnen sind, aus verschiedenen medizinischen Fachgebieten kommen und verschiedene Refundierungsmodelle vorweisen.

2.8 Anwendung der Assessment Frameworks

Methode FF4: Aussagen zu Anwendbarkeit der jeweiligen Frameworks
Einteilung in Risikoklassen nach MDR 2017/745
Besonderheiten und Limitationen in der Evaluation von DiGA

Zur Beantwortung der vierten Forschungsfrage wurden zu den ausgewählten Assessment Frameworks jeweils allgemeine Erläuterungen zur Anwendbarkeit des jeweiligen Frameworks formuliert (vgl. Kap. 3.1). In einem nächsten Schritt wurden die ausgewählten digitalen Gesundheitsanwendungen (vgl. Kap. 3.2 und 0) in die voraussichtliche Risikoklasse nach MDR 2017/745 eingeteilt und der derzeitigen Angabe einer Risikoklasse (laut Hersteller) gegenübergestellt. Außerdem wurden die digitalen Gesundheitsanwendungen in die jeweiligen Evidenzstufen nach dem NICE Evidence Standards Framework (vgl. Kap. 3.1.2) eingeordnet. Abschließend wurden die Assessment Frameworks auf die Evidenz der digitalen Gesundheitsanwendungen angewendet, um Besonderheiten und Limitationen der Evaluation digitaler Gesundheitsanwendungen zu beschreiben.

2.9 Auswahl der Länder

Methode FF5: Strategien zur DiGA-Refundierung von Österreich, Deutschland, Belgien, UK, Niederlande und Frankreich

Zur Beantwortung der fünften Forschungsfrage wurde eine Auswahl an Ländern getroffen, die bereits Herangehensweisen, Strategien und Modelle zur Refundierung digitaler Gesundheitsanwendungen vorweisen. Die Auswahl der Länder basierte auf dem MTRC Bericht „Reimbursement Landscape for Health Apps in Europe“ [24]. Folgende Länder wurden ausgewählt: Österreich, Deutschland, Belgien, UK, Niederlande und Frankreich. Die Informationen zu den verschiedenen Ansätzen der Refundierung digitaler Gesundheitsanwendungen der einzelnen Länder basieren zum einen auf dem MTRC Bericht und wurden zum anderen mit – in einer Handsuche gefundenen – Informationen in deutscher oder englischer Sprache ergänzt.

3 Ergebnisse

3.1 Übersicht Assessment Frameworks

Es liegt eine Vielzahl an Bewertungsmodellen für digitale Gesundheitsanwendungen vor. Allein in einer Literatursuche eingeschränkt ab 2018 wurden 554 Zitate gefunden: die meisten unter ihnen betreffen allerdings die konkrete Evaluation von spezifischen mHealth Anwendungen und nicht die Entwicklung eines generischen Instruments, das indikationsübergreifend eingesetzt werden kann.

Ein systematisches Review zu HTA Frameworks für eHealth Anwendungen (Suchzeitraum bis 2018) identifizierte und analysierte 21 Frameworks [13]. Dabei empfehlen die Autor*innen eine klare Definition der zu bewertenden Anwendung und schlagen hierfür die Template for Intervention Description and Replication (TIDieR) Checklist vor [25]. Als mögliches Bewertungsmodell für eHealth Anwendungen wird von den Autor*innen eine Kombination des TEMPEST Frameworks, des CHEATS Frameworks, des UVON Frameworks und des MAST Frameworks genannt. Ebenso wird das Evidence Standards Framework for Digital Health Technologies des NICE (National Institute for Health and Care Excellence) von den Autor*innen empfohlen [13].

Die Auswahl der hier beschriebenen Frameworks wurde in Bezug auf deren Aktualität, Ausgereiftheit und der zu erwartenden Anwendbarkeit auf digitale Gesundheitsanwendungen getroffen. Dennoch finden sich in der nachfolgenden Beschreibung Frameworks, die sich noch in Entwicklung und in Erprobung befinden.

Im folgenden Abschnitt werden die sechs ausgewählten Frameworks nach Jahr der Veröffentlichung, Zielgruppe, Methodik der Entwicklung sowie der Anwendung beschrieben und danach auf Ähnlichkeiten und Unterschiede analysiert. Die Kategorien der Datenextraktion der Assessment Frameworks sind in Kapitel 2.5 beschrieben.

Acht weitere Frameworks die für die detaillierte Analyse nicht herangezogen wurden sind im Kapitel 3.1.7 beschrieben. Die konkreten Ausschlussgründe der jeweiligen Frameworks werden ebenso im Kapitel 3.1.7 beschrieben.

In den ListServ Antworten der Mitglieder von INAHTA zeigte sich, dass sich viele Institutionen am NICE Evidence Standards Framework orientieren und dessen Nutzung (alleinstehend oder in Kombination) für eine erste Evaluierung von digitalen Gesundheitsanwendungen geplant ist. Darunter sind HIS (Schottland), HTW (UK), AquAS (Spanien), CADTH (Kanada) und IQWiG (Deutschland). Von AHTA (Australien) wurde ein HTA Module for Mobile Medical Applications entwickelt (siehe Kapitel 3.1.1), welches bisher noch nicht verwendet wurde. Interesse, dieses Framework zu nutzen, wurde von UVT (Italien) bekundet. Weitere Frameworks zur Bewertung von digitalen Gesundheitsanwendungen werden derzeit von AQuAS (Spanien) und INESSS (Kanada) entwickelt (vgl. Kapitel 7.2.1).

Einige HTA-Institutionen von INAHTA haben für die Bewertung digitaler Gesundheitsanwendungen bisher noch kein Assessment Framework genannt. Darunter sind SFOPH (Schweiz), AHTA (Australien), HAD (Uruguay), SBU (Schweden), SEC (Ukraine), HAS (Frankreich), CONITEC (Brasilien).

Vielzahl an indikationspezifischen Bewertungsmodellen, wenige generische Frameworks

SR von eHealth Bewertungsmodellen 2018 veröffentlicht

nach Aktualität, Reifegrad und Anwendbarkeit ausgewählt

im folgenden Detailbeschreibung von 6 Frameworks

Kurzbeschreibung zu 8 weiteren Frameworks

mehrere HTA-Institutionen orientieren sich an NICE Framework

weitere Frameworks in Entwicklung

einige HTA-Institutionen ohne Framework für DiGA

3.1.1 HTA Module for Mobile Medical Applications (MMAs) (AU)

Veröffentlichung April 2020	Das HTA Module for Mobile Medical Applications wurde im April 2020 von AHTA (Adelaide Health Technology Assessment) veröffentlicht.
Methodik der Entwicklung: SR & Interviews	Das Framework richtet sich an HTA-Expert*innen und Entscheidungsträger*innen. Dem Framework waren zwei Arbeiten zu Mobile Medical Applications (MMAs) vorausgegangen [12, 26]. Die erste Arbeit ist ein systematisches Review, in welchem gezeigt werden konnte, dass keines der identifizierten Frameworks alle Domänen eines HTAs abdeckt. Darüber hinaus wurden technologiespezifische Aspekte in den Frameworks nur unzureichend geprüft [12]. In der zweiten Arbeit wurden mögliche Risiken und regulatorische Aspekte von MMAs beleuchtet. Dabei konnte gezeigt werden, dass das indirekte Risiko durch Fehlinformation einer MMA nicht adäquat berücksichtigt wird [26].
Zielgruppe: HTA & Entscheidungsträger	Die Ergebnisse dieser Arbeiten wurden mit den Ergebnissen von neun semi-strukturierten Expert*inneninterviews (mit Therapeut*innen, Entwickler*innen und Entscheidungsträger*innen) zusammengetragen, um das HTA Module for Mobile Medical Applications zu entwickeln. Das Framework richtet sich an HTA Expert*Innen und Entscheidungsträger*innen.
Methode: klassische HTA Module (Checkliste) & technologie-spezifische Domänen	Das HTA Module for Mobile Medical Applications ist als technologiespezifische Checkliste zur Ergänzung klassischer HTA Domänen aufgebaut. Dabei werden für jede HTA-Domäne mögliche Herausforderungen der MMA aufgelistet. Diese sind beispielweise Software updates, Digital health literacy oder das Risiko durch Misinformation.
d. h. Haftungsfragen, Datenhoheit	In der legistischen Bewertungsdomäne befasst sich das HTA Module for Mobile Medical Applications mit Fragen zur Verantwortung und Haftung. Dabei gilt es zu klären, welche Partei (Mediziner*in, die eine MMA verschreibt, oder Hersteller der MMA) verantwortlich für die Korrektheit der Inhalte der MMA ist: Sind die von der MMA zur Verfügung gestellten Informationen aktuell und relevant? Entsprechen sie den Leitlinien und Empfehlungen von Fachgesellschaften? Außerdem muss festgelegt sein wer die Hoheit über die durch die MMA gesammelten Daten besitzt (Nutzer*in, Mediziner*in oder Dritte).
zusätzlich: zwei-zeitige Bewertung durch „Reappraisal“ post-market	Zusätzlich zu den klassischen HTA Domänen der klinischen Effektivität und der ökonomischen Evaluation sieht das HTA Module for Mobile Medical Applications ein „Reappraisal“ (Neubewertung) der digitalen Gesundheitsanwendung vor. In diesem „Reappraisal“ sollen die Post-Market-Surveillance, insbesondere durch Software Updates hervorgerufene neue Funktionen der digitalen Gesundheitsanwendung, bewertet werden.
soziale Domäne nicht obligatorisch	Im HTA Module for Mobile Medical Applications sind alle Bewertungsdomänen (bis auf die soziale Bewertungsdomäne) verpflichtend zu bewerten [27]. Details zum HTA Module for Mobile Medical Applications Framework finden sich im Anhang, Kapitel 7.1.1.

3.1.2 Evidence Standards Framework for Digital Health Technologies (ENG)

Das Evidence Standards Framework for Digital Health Technologies wurde im März 2019 veröffentlicht. Es entstand aus einer Kooperation von NICE (National Institute for Health and Care Excellence) mit dem NHS (National Health Service) England, Public Health England und MedCity. Das Framework richtet sich sowohl an Hersteller von digitalen Technologien als auch an Entscheidungsträger*innen. Die Methodik der Entwicklung wird in den Publikationen zum Framework nicht beschrieben.

Das Framework gliedert sich grob in zwei Sektionen. Sektion A befasst sich mit Standards zur geforderten Evidenz zur Effektivität von digitalen Gesundheitsanwendungen. Sektion B befasst sich mit Standards zur geforderten Evidenz zu ökonomischen Auswirkungen von digitalen Gesundheitsanwendungen. Das besondere Merkmal des NICE Evidence Standards Framework ist, die geforderte Evidenz zu Effektivität in Relation zum möglichen Risiko (das von der digitalen Gesundheitsanwendung ausgeht) zu setzen. Es erfolgt eine Stufeneinteilung der digitalen Gesundheitsanwendungen, eine höhere Stufe birgt potentiell größere Risiken und erfordert ein höheres Maß an Evidenz. Dabei werden vier Stufen unterschieden:

- *Stufe 1:* Digitale Gesundheitsanwendungen ohne individuell messbare Effekte auf die Gesundheit von Anwender*innen, organisatorische Anwendungen auf Systemebene. Beispiele: elektronische Gesundheitsakte, Krankenhaus-Informationssysteme.
- *Stufe 2:* Digitale Gesundheitsanwendungen, welche Gesundheitsinformationen anbieten, einfache Monitoringfunktionen übernehmen oder Anwendungen zur Kommunikation. Beispiele: Anwendungen mit Empfehlungen für einen gesunden Lebensstil, digitaler Kopfschmerzkalender, Anwendungen für Video-Sprechstunden mit Therapeut*innen.
- *Stufe 3a:* 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. Beispiele: Anwendungen zur Raucherentwöhnung oder Gewichtsreduktion.
- *Stufe 3b:* 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. Beispiele: 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.

Das Framework ist auch auf digitale Gesundheitsanwendungen mit Artificial Intelligence (AI) mit fixierten Algorithmen anwendbar. Von diesem Framework explizit ausgenommen sind jedoch digitale Gesundheitsanwendungen die AI mit adaptiven Algorithmen (machine learning, selbstlernende Algorithmen, die ihr Verhalten basierend auf verfügbaren Informationen kontinuierlich und automatisiert anpassen) verwenden [28].

Veröffentlichung
März 2019

Zielgruppe:
Entscheidungsträger
& Hersteller

Methodik der Anwendung:
Systematik zu Evidenz zur
Effektivität und zu GesÖk

Evidenz zu Effektivität
in Relation zu Risiko der
mHealth Anwendung

Systematik
Risikostufe 1:
organisatorische
Anwendungen

Risikostufe 2:
Monitoring und
Kommunikation

Risikostufe 3a:
individuelle Messungen

Risikostufe 3b:
individuelle Diagnostik
und Therapie

auch auf AI anwendbar:
NUR bei fixierten
Algorithmen

adaptive Algorithmen
explizit ausgenommen!

**IQWiG:
Adaption des
NICE-Frameworks**

Das NICE-Framework bekommt international große Aufmerksamkeit. So empfiehlt das IQWiG in den INAHTA ListServ Antworten die Anwendung des NICE-Frameworks (vgl. Appendix 7.2.1). Zudem wurde in einer Publikation des Netzwerks evidenzbasierter Medizin eine Visualisierung der NICE Framework Anforderungen an neue digitale Gesundheitsanwendungen erstellt (vgl. Appendix 7.1.2 [29]).

3.1.3 Digi-HTA (FIN)

**Veröffentlichung
November 2019**

**Methodik der Entwicklung:
SR & Interviews**

**Zielgruppe:
HTA & Entscheidungsträger**

Im November 2019 wurde von der Universität Oulu (Finnland) ein Framework zur Bewertung von digitalen Gesundheitsanwendungen veröffentlicht.

Das sogenannte „Digi-HTA“ Framework wurde nach einer systematischen Literatursuche zu Bewertungen von mHealth, AI und Robotik entwickelt. Ergänzend dazu wurden Expert*inneninterviews mit Herstellern der Technologien und Anbieter*innen von Gesundheitsdienstleistungen durchgeführt. Außerdem wurden während der Entwicklung des Digi-HTA Frameworks vier multiprofessionelle Workshops (mit HTA-Expert*innen, AI-Expert*innen und Mediziner*innen) durchgeführt. Die Zielgruppe des Digi-HTA Frameworks sind HTA-Expert*innen und Entscheidungsträger*innen.

**Methode: klassische
HTA Module (Fragebogen)
& technologie-spezifische
Domänen**

Das Digi-HTA Framework ist als Fragebogen aufgebaut. Dabei gibt es zu den verschiedenen HTA-Domänen (technische Aspekte, Effektivität, Sicherheit, ökonomische Aspekte, Anwenderaspekte und Zugang, organisatorische Aspekte) qualitative Fragestellungen. Spezielle Domänen des Frameworks sind Interoperabilität, Künstliche Intelligenz und Robotik (mit jeweils domänenspezifischen Fragestellungen).

**d. h. Interoperabilität,
AI, Robotik;
NICHT im Assessment,
erst im Appraisal:
Datenschutz
und -sicherheit**

Die Themen Datensicherheit und Datenschutz werden von diesem Framework nicht bewertet, da deren Prüfung ausgelagert und in den sogenannten „Data Security and Protection Preliminary Task“ und „Information Security and Data Protection Requirements“ von Datenschutzexpert*innen vorgesehen ist. Die Bewertung von Datensicherheit und Datenschutz der Datenschutzexpert*innen fließt in die anschließende HTA-Empfehlung für oder gegen ein Produkt mit ein.

**rechtliche, soziale und
ethische Aspekte werden
ausgeschlossen**

Das Digi-HTA Framework ist für eine rasche Bewertung von sich schnell entwickelnden Technologien entwickelt worden. Das Digi-HTA ist das erste Framework, welches die Themen mHealth, AI und Robotik in einem Framework vereint. Dabei klammert das Digi-HTA Framework rechtliche, soziale und ethische Aspekte bewusst aus, da diese in ihrer Bewertung komplex und zeitintensiv sind [30].

Details zum Digi-HTA Framework finden sich im Anhang, Kapitel 7.1.3.

3.1.4 Digital Health Interventions (DEDHI) Framework (CH)

**Veröffentlichung
November 2019**

**Zielgruppe:
Entwickler & Evaluatoren**

Im November 2019 wurde vom Center for Digital Health Interventions (eine Initiative der Universität St. Gallen & ETH Zürich) das Design and Evaluation Framework for Digital Health Interventions (DEDHI Framework) veröffentlicht. Das DEDHI Framework richtet sich an Wissenschaftler*innen unterschiedlicher Disziplinen (Verhaltensmedizin, Medizinische Informatik) und Evaluator*innen.

Das DEDHI Framework ist als Checkliste zu Forschung & Entwicklung sowie Implementierung von Digital Health Interventions (DHIs) aufgebaut. Das DEDHI Framework basiert auf einer adaptierten Version der dreiphasigen „Multiphase optimization strategy“ (MOST) [31] und erweitert diese um eine vierte Phase. Die drei Phasen der MOST sind Vorbereitungsphase, Optimierungsphase und Bewertungsphase. Die durch das DEDHI Framework erweiterte vierte Phase beschreibt die Phase der Implementierung und deren Ziele, Aufgaben, technische Ausgereiftheit, Bewertungskriterien und zu überwindende Barrieren. Mögliche Barrieren der Implementierungsphase sind Entscheidungen zur Refundierung, Langzeitfinanzierung und Maintenance-Kosten. Vor Erstellung des DEDHI Frameworks wurde eine systematische Literatursuche zur Bewertung von digitalen Gesundheitsinterventionen durchgeführt. Dabei wurden 331 Bewertungskriterien und 98 Barrieren in der Implementierung identifiziert, konsolidiert und den vier Produktentwicklungsphasen zugeordnet.

Die geforderte Evidenz richtet sich nach Produktentwicklungsstadium: In der ersten Phase Studien zu Durchführbarkeit; in der zweiten Phase Studien zu Optimierung; in der dritten Phase RCTs; in der vierten Phase werden kontinuierliche Erhebungen zu Reichweite, Impact und Nebenwirkungen vorgeschlagen [32].

Details zum DEDHI Framework finden sich im Anhang, Kapitel 7.1.4.

Methodik der Entwicklung:
Technologie-
entwicklungsphasen
(MOST) als Basis &
SR zu Bewertungskriterien

**Zuordnung der Kriterien
nach 4 Phasen**

Methodik der Anwendung:
Evidenzerfordernisse nach
Phase

3.1.5 Digital Health Scorecard (USA)

Die Digital Health Scorecard wurde im Mai 2019 durch das Armstrong Institute for Patient Safety and Quality & Johns Hopkins Medicine (Baltimore, MD, USA) veröffentlicht. Die Digital Health Scorecard richtet sich an unabhängige Evaluator*innen und soll jeweils einmal vor als auch nach Marktzulassung angewendet werden. Die Methodik der Entwicklung wird in den Publikationen zum Framework nicht beschrieben.

Die Digital Health Scorecard besteht aus vier Komponenten: Technische Aspekte, klinische Effektivität (Nutzen), Benutzerfreundlichkeit und Kosten. Zu jeder dieser Komponenten werden technologiespezifische Betrachtungen und Erwägungen aufgelistet. Die geforderte Evidenz zur klinischen Effektivität sieht einen Vergleich zum existierenden Goldstandard auf zuvor definierte klinische Endpunkte vor. Die Digital Health Scorecard macht keine Angaben zu vorgeschlagenen Studiendesigns. Die Ergebnisse der Bewertungen der vier Komponenten sollen in einem Global Digital Health Score zusammengetragen werden. Die Digital Health Scorecard befindet sich noch in Entwicklung – Scores wurden bisher noch nicht definiert [33].

Details zur Digital Health Scorecard finden sich im Anhang, Kapitel 7.1.5.

Veröffentlichung
Mai 2019

Zielgruppe: Evaluatoren

Methodik der Anwendung:
Bewertung in
4 Domänen,
Evidenz zu Vergleich
mit Goldstandard

**Ergebnis: Global Score
(erst in Entwicklung)**

3.1.6 mHealth Agile Development & Lifecycle (CA)

Veröffentlichung	Im September 2018 wurde von der Universität Ottawa (Kanada) das mHealth Agile Development & Lifecycle Framework veröffentlicht. Das mHealth Agile Development & Lifecycle Framework richtet sich an Technologieentwickler*innen und Entscheidungsträger*innen.
September 2018,	
Zielgruppe:	
Entwickler &	
Entscheidungsträger	
Methodik der Entwicklung:	Das mHealth Agile Development & Lifecycle Framework orientiert sich am IDEAS Framework (Integrate, Design, Assess and Share Framework, vgl. Kapitel 3.1.7) und ist als Fragebogen mit kontextuellen Fragen aufgebaut. Nach einer Phase der Projektidentifikation (Identifikation eines Problems, Lösungskonzeption, Anwenderforschung und Fragen zur Finanzierung) folgen 4 Phasen zur klinischen Evaluation und zu Evidenzerfordernissen je nach Produktentwicklungsstadium:
Technologieentwicklungsphasen (IDEAS) als Basis	
Methoden der Anwendung:	■ <i>Phase 1:</i> User-experience Design, Development & Alpha-Testing ■ <i>Phase 2:</i> Beta Testing (Kohortenstudien, Fallserien) ■ <i>Phase 3:</i> Clinical Trial Evaluation (RCTs, A/B Testing) ■ <i>Phase 4:</i> Post-Market Surveillance
Evidenzerfordernisse nach Phase	
Berücksichtigung der agilen Softwareentwicklung	Bei der Entwicklung einer digitalen Gesundheitsanwendung werden häufig Methoden der agilen Softwareentwicklung angewendet. Bei der agilen Softwareentwicklung wird durch kontinuierliches und iteratives Feedback der Nutzer*innen an die Entwickler*innen eine erhöhte Flexibilität und Transparenz in der Entwicklung von Technologien erzielt. Der Ansatz des mHealth Agile Development & Lifecycle Frameworks ist es, diese agile Softwareentwicklung mit den derzeit angewandten Methoden zur Bewertung herkömmlicher Gesundheitstechnologien zu kombinieren [34].
	Details zum mHealth Agile Development & Lifecycle Framework finden sich im Anhang, Kapitel 7.1.6.

3.1.7 Weitere Frameworks

zahllose weitere Frameworks	Es liegt eine Vielzahl weiterer Frameworks vor. Auf diese weiteren Frameworks wird in diesem Bericht nicht detailliert eingegangen, da sie entweder einen zu engen „Scope“ (z. B. nur Evaluation der Benutzerfreundlichkeit: MAUQ; nur zum Vergleich von mHealth Anwendungen untereinander: MARS) haben oder die Aktualität des Frameworks in Bezug auf mHealth nicht ausreichend gegeben ist (z. B. MAST, CHEATS). Dennoch sollen sie im folgenden Abschnitt eine kurze Erwähnung finden.
Telemedizin enger Scope	
CHEATS 2002: ICT	■ Das „<i>Generic Information Communication Technology Evaluation</i>“ <i>Framework (CHEATS) (2002)</i>: Das CHEATS Framework dient der Bewertung von Informations- und Kommunikationstechnologien. Es wurde für Telepsychiatrie, Teledermatologie und Telebildung eingesetzt [35]. Aufgrund mangelnder Aktualität wurde es von der Detailanalyse ausgeschlossen.
MAST 2012: Telemedizin	■ Das „<i>Model for Assessment of Telemedicine Applications</i>“ <i>(MAST) (2012)</i>: Das MAST Framework basiert auf dem EUnetHTA Core Model und den dort entwickelten neun Domänen. Es ist zur Bewertung von ausgereiften Anwendungen im Bereich der Telemedizin vorgesehen und an zahlreichen Projekten auch bereits erprobt worden [36]. Aufgrund mangelndem Bezug zu mHealth wurde es von der Detailanalyse ausgeschlossen.

- Das „*Technology, Economic, Market, Political, Evaluation, Social and Transformation*“ *Framework (TEMPEST)* (2012): Das TEMPEST Framework dient der Bewertung von Informations- und Kommunikationstechnologien, insbesondere deren Einführung und Umsetzung [37]. Aufgrund mangelndem Bezug zu digitalen Gesundheitsanwendungen wurde es von der Detailanalyse ausgeschlossen.
- Die „*Mobile App Rating Scale*“ (MARS) (2015): Die MARS Skala dient dem Vergleich von digitalen Gesundheitsanwendungen untereinander. In die Bewertung fließen Engagement, Funktionalität, Ästhetik und Informationsqualität ein [38]. Aufgrund des zu engem „Scopes“ (nur Vergleich von digitalen Gesundheitsanwendungen untereinander) wurde es von der Detailanalyse ausgeschlossen.
- Das „*Integrate, Design, Assess and Share*“ *Framework (IDEAS)* (2016): Das IDEAS Framework dient zur Anleitung der Entwicklung und Bewertung effektiverer digitaler Interventionen. IDEAS ist in vier Entwicklungsphasen gruppiert (Integrate, Design, Assess, and Share) und zielt auf die Beschreibung der notwendigen Datengenerierung resp. Evidenz in der jeweiligen Phase ab und soll zur Unterstützung von Entwicklern dienen [39]. Da die Bewertung digitaler Interventionen nur ein Teilaspekt dieses Frameworks ist, wurde es von der Detailanalyse ausgeschlossen.
- Die „*Unified eValuation using Ontology*“ (UVON) (2016): Das UVON Framework basiert auf dem FI-STAR (Future Internet Social and Technological Alignment Research) und dem MAST Framework. Mittels Vereinheitlichung und Aggregation der Qualitätsmerkmale können Informationstechnologien im Gesundheitswesen bewertet werden [40]. Aufgrund mangelnder Aktualität in Bezug auf technologiespezifische Aspekte wurde es von der Detailanalyse ausgeschlossen.
- Das „*Organization for the Review of Care and Health Applications*“ *Framework (ORCHA)* (2017): Das ORCHA Framework besteht aus 24 Kriterien zur Evaluation von digitalen Gesundheitsanwendungen mit Fokus auf Datenschutz, Benutzerfreundlichkeit und Partizipation [41]. Aufgrund mangelnder Aktualität in Bezug auf technologiespezifische Aspekte wurde es von der Detailanalyse ausgeschlossen.
- Der „*mHealth App Usability Questionnaire*“ (MAUQ) (2019): MAUQ ist ein Fragebogen um die Benutzerfreundlichkeit von digitalen Gesundheitsanwendungen zu bewerten, bestehend aus 21 Fragen zu Zufriedenheit, Anordnung von Information und Nützlichkeit [42]. Da sich dieses Framework allein der Bewertung der Benutzerfreundlichkeit widmet, wurde es von der Detailanalyse ausgeschlossen.

TEMPEST 2012:
ICT

MARS 2015:
mHealth Vergleiche

IDEAS 2016:
Datengenerierung in Entwicklungsphasen

UVON 2016:
Weiterentwicklung MAST

ORCHA 2017:
mHealth Evaluation

MAUQ 2019:
Benutzerfreundlichkeit

Tabelle 3.1-1: Ausgewählte Frameworks zur Evaluation von mHealth Anwendungen

	HTA Module for Mobile Medical Applications [27]	Evidence Standards Framework for Digital Health Technologies [28]	Digital Health Care Services/Digi-HTA [30]	Design & Evaluation Framework for Digital Interventions/ DEDHI [32]	Digital Health Scorecard [33]	mHealth Agile Development & Lifecycle [34]
Organisation/ Universität/ Land	AHTA Adelaide Health Technology Assessment University of Adelaide, Australia	NICE National Institute for Health and Care Excellence, United Kingdom	Faculty of Medicine, University of Oulu, Finland	Center for Digital Health Interventions, University of St. Gallen & ETH Zürich, Switzerland	Armstrong Institute for Patient Safety and Quality & Johns Hopkins Medicine Baltimore, MD, USA	Faculty of Medicine, University of Ottawa, Canada
Finanzierung	Australian Government Research Training Program Scholarship and University of Adelaide, School of Public Health	National Health Service England	Regional Council of Northern Ostrobothnia (funding from the European Regional Development Fund), Funding by the Northern Ostrobothnia Hospital District.	Health Promotion Switzerland, and the European Social Fund and the Free State of Saxony (Grant no. 100310385).	Das Framework entstand ohne Finanzierung aus dem öffentlichen, kommerziellen oder gemeinnützigen Sektor.	Public Health Agency of Canada, Sanofi-Pasteur, the Canadian Institutes of Health Research, the Ottawa Hospital and Health Canada.
Jahr	April 2020	März 2019	November 2019	November 2019	Mai 2019	September 2018
Zielgruppe	HTA Expert*innen, Entscheidungsträger*Innen	Technologieentwickler*Innen, Entscheidungsträger*Innen	HTA Expert*innen, Entscheidungsträger*Innen	Wissenschaftler*Innen, Evaluator*Innen	Unabhängige Evaluator*Innen	Technologieentwickler*Innen, Entscheidungsträger*Innen
Status Framework	in Entwicklung	Etabliert, in Erprobung	in Entwicklung/in Erprobung	in Erprobung	in Entwicklung	in Erprobung
Anwendungsbereich/Titel	Mobile Medical Application (MMA)	Digital Health Technologies (DHT)	mHealth, KI, Robotik	Digital Health Interventions (DHI)	Digital Health Products	mHealth/ Digital Health Products
Risikoklassen	Keine Erwähnung von Risikoklassen	Risiko Klassifikation (1, 2, 3a/3b) entsprechend der Funktionalität der DHT	Keine definierten Risikoklassen, aber deskriptive Fragen nach klinischer Konsequenz	Keine Erwähnung von Risikoklassen	Keine definierten Risikoklassen, aber Differenzierung nach Funktionalitäten und deren klinischer Konsequenz	Keine definierten Risikoklassen, aber Differenzierung zwischen niedrig Risiko (keinerlei Risiken für Anwender*Innen) und Hochrisiko Anwendungen
Evidenzerfordernisse	Keine Erwähnung von Studiendesigns, nur von Vergleichsinterventionen, Analyse von Schaden durch Fehlinformation, ethische Aspekte	Evidenzerfordernisse entsprechend der Risikoklasse (RK), je RK Minimalerfordernisse vs. Best Practice, z. B. 3a: Beobachtungsstudie vs. Interventionsstudie, 3b: Interventionsstudien/RCT Komparatoren, relevante Endpunkte	Empfohlene Studiendesigns: Fallstudien, RCTs, Systematische Reviews	Orientierung am MOST Framework zu Evidenzerfordernissen in Produktentwicklungsphasen Phase 1: Studien zu Durchführbarkeit Phase 2: Studien zu Optimierung Phase 3: RCTs Phase 4: kontinuierliche Erhebungen zu Reichweite, Impact und Nebenwirkungen	Keine Erwähnung von Studiendesigns, nur kritische Bewertung der Evidenz bezgl. Impact auf definierte klinische Endpunkte. Vergleiche zu Gold Standard.	Orientierung am IDEAS Framework zu Evidenzerfordernissen in Produktentwicklungsstadien Phase 2: Kohortenstudien, Fallserien Phase 3: RCTs (sofern mHealth Produkte als Medizinprodukte definiert sind), A/B testing (Testung von App Versionen)
Zeitpunkt der Anwendung des Frameworks/Intention	vor Refundierung	vor Refundierung zur Evidenz-Generierung	vor Refundierung zur Evidenz-Generierung, Post-market surveillance	Zur Evidenz-Generierung, Post-market surveillance	Zur Evidenz-Generierung, Post-market surveillance	zur Evidenz-Generierung, Post-market surveillance

	HTA Module for Mobile Medical Applications [27]	Evidence Standards Framework for Digital Health Technologies [28]	Digital Health Care Services/Digi-HTA [30]	Design & Evaluation Framework for Digital Interventions/ DEDHI [32]	Digital Health Scorecard [33]	mHealth Agile Development & Lifecycle [34]
Format des Frameworks	herkömmliche HTA-Checkliste	Fragen zum Kontext zur Einordnung der DHT in eine Risikoklasse, tabellarische Evidenzerfordernisse (kumulativ, aufsteigend)	Fragebogen HTA-Checkliste	Checkliste zu F&E sowie Implementierung von DHLs	Scorecard (Scores sind nicht definiert)	Fragebogen zu F&E von mHealth Apps im Lebenszyklus nach IDEAS
Bewertungsdomänen	Technische Aspekte, Effektivität (Nutzen), Sicherheit Kosten, Kosteneffektivität, organisatorische, legistische, ethische und soziale Aspekte	1: Plausibilität, Relevanz, Akzeptanz, Gleicher Zugang, Genauigkeit. 2: Verlässlichkeit, Wert für Zielpopulation, Qualität und Datensicherheit. 3a/3b: klinische Effektivität anhand relevanter Endpunkte, 1-3: ökonomischer Impact	Technische Aspekte Effektivität (Nutzen), Sicherheit, ökonomische Aspekte, Anwenderaspekte und Zugang, organisatorische (Interoperabilität)	Benutzerfreundlichkeit, inhaltliche Qualität, Privatsphäre und Datenschutz, Verantwortlichkeit, Adhärenz, Ästhetik, empfundener Nutzen, Effektivität (Nutzen), Qualität des Service, Personalisierung, empfundenes Engagement, ethische Aspekte, Sicherheit	4 Domänen: Technische Aspekte, klinische Effektivität (Nutzen), Benutzerfreundlichkeit, Kosten	Phase 1 der Entwicklung: Technische Aspekte, Performanz und Datensicherheit Phase 2: Anwenderaspekte, Akzeptanz Phase 3: Nutzen
Technologie-spezifische Aspekte: Datenschutz Datensicherheit Umgang mit Updates Künstliche Intelligenz (KI)	Datenschutz: ja Datensicherheit: ja Umgang mit Updates: ja Künstliche Intelligenz: k.A.	Datenschutz: k.A Datensicherheit: k.A Umgang mit Updates: k.A Künstliche Intelligenz: nur fixe Algorithmen, NICHT adaptive Algorithmen	Datenschutz: ja Datensicherheit: ja Umgang mit Updates: ja Künstliche Intelligenz: ja	Datenschutz: ja Datensicherheit: ja Umgang mit Updates: ja Künstliche Intelligenz: nein	Datenschutz: ja Datensicherheit: ja Umgang mit Updates: nein Künstliche Intelligenz: nein	Datenschutz: ja Datensicherheit: ja Umgang mit Updates: ja Künstliche Intelligenz: nein
Involvierung Patient*Innenvertretung	k.A.	k.A.	k.A.	nein	nein	k.A.
Framework-Spezifika Resümee	Systematischer Review zu derzeit verwendeten Evaluations-Frameworks. Derzeit kein eigenständiges Framework, nur Zusammenschau. Traditionelle HTA-Perspektive.	Stellt die geforderte Evidenz in Relation zum möglichen Risiko der DHT. Auflistung geforderter Studiendesigns je nach Risikoklasse. Einziges hoch-entwickeltes Framework zur sofortigen Anwendung.	Framework basiert auf Systematischem Review, Interviews und Workshop. Vereint mHealth, KI und Robotik in einem Framework. Ethische, soziale und rechtliche Aspekte sind explizit kein Bestandteil dieses Frameworks.	Framework behandelt unterschiedliche Entwicklungsphasen einer Digital Health Intervention, die Implementierungsphase wird detailliert berücksichtigt.	Die Digital Health Scorecard soll jeweils einmal vor als auch nach Marktzulassung angewendet werden.	Das Framework vereint einen „agilen“ Entwicklungsprozess aus Perspektive von Technologienentwickler*Innen mit Methoden zur Nutzenbewertung.

3.1.8 Zusammenfassung und Diskussion der ausgewählten Frameworks

generische Frameworks zur indikationsübergreifenden Bewertung ausgewählt	Die hier beschriebenen Frameworks können für digitale Gesundheitsanwendungen aller medizinischen Fachgebiete und Populationen eingesetzt werden. Zusätzlich existieren Frameworks, die sich speziell einem Fachgebiet, einer Indikation oder einer Population widmen. Beispiele hierfür sind das App Evaluation Model der American Psychiatric Association [43], oder ein Framework für die Bewertung von digitalen Gesundheitsanwendungen für Jugendliche [44]. Die Frameworks bedienen sich auch unterschiedlicher Bezeichnungen für ihren Anwendungsbereich („Mobile Medical Applications“, „Digital Health Technologies“, „Digital Health Interventions“, „Digital Health Products“), was den Mangel an exakten Definitionen (vgl. Kapitel 1) widerspiegelt.
Bezeichnungen der Frameworks: Digital Health, mHealth	Die beschriebenen Frameworks basieren zum Teil auf systematischen Literaturanalysen zu Bewertungskriterien und bestehenden Frameworks zur Bewertung von nicht-digitalen Technologien (HTA-Domänen), sowie bestehenden Technologie-Entwicklungsmodellen (MOST, IDEAS) und wurden – zum Teil – interdisziplinär von verschiedenen Interessensgruppen konzipiert. So haben Technologie-Entwickler*innen, Mediziner*innen, Therapeut*innen, Entscheidungsträger*innen und Regulator*innen ihre Beiträge zu den Frameworks geleistet. Die Patient*innenvertretung scheint in der Konzeption der Frameworks keine oder nur eine geringe Rolle gespielt zu haben. Wenig ist zu den Entwicklungsschritten des NICE-Frameworks bekannt, es dürfte sich aber an der Medizinprodukteverordnung und den dortigen Risikoeinstufungen orientiert haben.
Methoden der Entwicklung: SR, Adaption bestehender Frameworks, Einbindung von Stakeholdern ohne Patient*innen NICE: orientiert an Risikoeinstufung von MDR/IVDR	Keines der beschriebenen Frameworks ist vollständig etabliert und in der Anwendung getestet und indikationsübergreifend an konkreten Beispielen validiert. Die beschriebenen Frameworks unterscheiden sich nach Format (Checkliste, Fragebogen, Tabelle), Spektrum der Bewertungsdomänen, technologiespezifischen Aspekten und Zielgruppe (vgl. Tabelle 1.3-1).
bislang keine Etablierung Format der Anwendung: Checklisten & Fragebögen Risikoklassen werden nur bei NICE berücksichtigt: Voraussetzung für Entscheidungen zu Bedarf nach Art & Umfang von Evidenz bereits 2016 von IGES empfohlen	Auffallend ist, dass die meisten Frameworks das Risiko fehlerhafter Anwendungen mit entsprechenden klinischen Konsequenzen zwar berücksichtigen, jedoch keine Einteilung nach definierten Risikoklassen vorschlagen. Eine präzise Einteilung nach definierten Risikoklassen bietet nur das Evidence Standards Framework von NICE, was aber eine wesentliche Voraussetzung ist, um in der Vielzahl unterschiedlicher mHealth Anwendungen Unterscheidungen im Bedarf nach Art und Umfang von Evidenz zu treffen. Diese Herangehensweise für eine Regulierung nach Risikoklassen von digitalen Gesundheitsanwendungen wurde bereits 2016 durch das IGES Institut empfohlen. Das Risiko hängt dabei stark von der Art der erhobenen Daten und deren Verarbeitung ab [10].
abgestuftes Vorgehen nach Entwicklungs- und Risikostufen erleichtert die Bewältigung der Vielzahl der Bewertungen	Nicht in allen beschriebenen Frameworks ist also ein abgestuftes Vorgehen im Assessment vorgesehen. Ein abgestuftes Vorgehen im Assessment von digitalen Gesundheitsanwendungen kann jedoch zeitsparend sein. So besteht beispielsweise bei Anwendung eines stufenhaften Frameworks noch kein Bedarf die klinische Evidenz zu bewerten, wenn die digitale Gesundheitsanwendung die Kriterien für Datenschutz und Datensicherheit nicht adäquat erfüllt [45]. Ein Beispiel für eine solche stufenhafte Bewertung ist das mHealth Agile Development & Evaluation Lifecycle Framework [34].

Die Frameworks unterscheiden sich auch hinsichtlich des vorgeschlagenen Zeitpunkts des Assessments. So ist das NICE Evidence Standards Framework für eine Bewertung ausgereifter digitaler Gesundheitsanwendungen vorgesehen [28]. Demgegenüber steht die Empfehlung, eine begleitende Bewertung bereits in allen Phasen einer digitalen Gesundheitsanwendung durchzuführen [46].

Die Frameworks unterscheiden sich in Bezug auf die Abdeckung aller Bewertungsdomänen eines klassischen HTAs (vgl. Tabelle 1.3-1). An dieser Stelle hervorzuheben ist, dass das Digi-HTA Framework die Bewertung ethischer, sozialer und legisistischer Fragen aufgrund von Komplexität und Zeitintensität ausblendet [30]. Auch im MMA-Framework werden die sozialen Aspekte als weniger relevant erachtet [27].

Nicht alle technologie-spezifischen Aspekte (wie Datenschutz, Datensicherheit, Umgang mit Updates, Aspekte der künstlichen Intelligenz) finden ausreichend Berücksichtigung in den Assessment Frameworks. So werden die Bewertung von Datensicherheit und Datenschutz nicht als ein integrativer Teil der Evaluation, sondern erst als ein Aspekt des Appraisals und der Empfehlung beschrieben (Digi-HTA, [30]).

Aspekte der künstlichen Intelligenz (AI) insbesondere zur Transparenz der verwendeten Logik der Algorithmen, findet in den Assessment Frameworks nur wenig Beachtung. Die Autor*innen des Digi-HTA Frameworks betonen die Notwendigkeit von AI Transparenz, insbesondere wenn die Empfehlung einer A.I. sich nicht mit der Empfehlung einer Therapeut*in deckt [30].

Das HTA Module for Mobile Medical Applications und die Digital Health Scorecard geben keine Erwähnung von geforderten Studiendesigns an [27, 33]. Das NICE Evidence Standards Framework for Digital Health Interventions, das Digi-HTA, das DEDHI Framework und das mHealth Agile Development & Evaluation Lifecycle Framework nennen insbesondere RCTs als das gewünschte Studiendesign.

Darüber hinaus könnten neuartige Studiendesigns (vgl. Kapitel 4.2.4) in der Bewertung von digitalen Gesundheitsanwendungen geeignet sein [28, 30, 32, 34]. In einem systematischen Review von Kolasa et al. wird beschrieben, dass der Ansatz der Messung der qualitätsadjustierten Lebensjahre (QALYs) für digitale Gesundheitsanwendungen möglicherweise nicht geeignet ist [47].

Eine Übersicht von Bewertungskriterien von Digital Health Technologies und deren Kategorisierung hat gezeigt, dass Aspekte der Benutzerfreundlichkeit am häufigsten bewertet werden, während ethische Aspekte und Sicherheit von Digital Health Technologies bisher am wenigsten in die Bewertung von Digital Health Technologies einfließen [32]. Zu den Kosten und der Kostenwirksamkeit von digitalen Gesundheitsanwendungen wurde bisher kaum Evidenz gesammelt [48].

Zeitpunkt der Bewertung wesentlich: in Entwicklung, vor Refundierung Refundierung mit Begleitforschung

Unterschiede im Spektrum der zu bewertenden HTA-Aspekte

Unterschiede bei Berücksichtigung von Technologie-spezifischen Aspekten

Transparenz bei AI: Logik des zugrundeliegenden Algorithmus

Studiendesigns: Wunsch in hohen Risikoklassen sind RCTs

neuartige Studiendesigns von zunehmender Relevanz

Empirie der Bewertung von mHealth Anwendungen: Benutzerfreundlichkeit oft, ethische Aspekte, Datensicherheit, Kosten selten bewertet

3.2 Refundierte digitale Gesundheitsanwendungen in ausgewählten Ländern

DiGA mit Selektivverträgen

Einteilung nach Anwendungsgebiet, Input, Output, Kosten, Refundierungsmodell

Im folgenden Abschnitt werden digitale Gesundheitsanwendungen beschrieben, die in ausgewählten Ländern von einzelnen Sozial-/Krankenversicherungen bereits refundiert werden oder eine Refundierung beantragt wurde. Nach einer Kurzbeschreibung der Funktionsweise der digitalen Gesundheitsanwendungen folgt eine Einteilung nach Anwendungsgebieten, verwendetem Input (Messwerte, Text, Bild, Ton) und Output (informativ, diagnostisch, therapeutisch), sowie eine Analyse der Kosten bei Verwendung und vorhandene Refundierungsmodelle.

**Handsuche:
38 DiGA identifiziert,
Detailanalyse von
11 DiGA**

Im Rahmen der Handsuche nach digitalen Gesundheitsanwendungen wurden 38 digitale Gesundheitsanwendungen identifiziert. Aus diesen wurden 11 für eine Detailanalyse ausgewählt. Die Auswahlkriterien für die analysierten digitalen Gesundheitsanwendungen ist im methodischen Vorgehen (Kapitel 2.7) beschrieben. Diese Auswahl ist jedenfalls nur als Illustration zu erachten und keineswegs vollständig, da das Gebiet der digitalen Gesundheitsanwendungen ein sehr dynamisches ist und eine umfassende Dokumentation unmöglich erscheint.

3.2.1 Übersicht zu und Kurzbeschreibung von ausgewählten digitalen Gesundheitsanwendungen

Kurzbeschreibung ausgewählter DiGA

Im Folgenden werden die digitalen Gesundheitsanwendungen kurz beschrieben. Die Tabelle 3.2-1 gibt eine Übersicht über die Anbieter, Funktionsweisen (und die verwendeten Informationen sowie deren Verarbeitung) und Zielgruppen der identifizierten und ausgewählten digitalen Gesundheitsanwendungen. Die Einteilung dieser digitalen Gesundheitsanwendungen in die entsprechenden Risikoklassen und die identifizierte Evidenz aus Studien findet sich im Kapitel 0.

SkinVision (SkinVision B.V.):

- SkinVision ist eine digitale Gesundheitsanwendung die der Früherkennung von Hautkrebs dient. Mittels Smartphonekamera werden Bildaufnahmen von Muttermalen (Nävi) aufgenommen. Durch einen Algorithmus können Anzeichen von Hautkrebs (Melanom, Basalzellkarzinom, Plattenepithelkarzinom) berechnet und erkannt werden.

Ada (Ada Health GmbH):

- Ada ist eine digitale Gesundheitsanwendung, die dem/der Anwender*in Fragen zu Symptomen stellt. Die Antworten werden mit Hilfe eines medizinischen Wörterbuchs ausgewertet und mit Diagnosen abgeglichen. Anhand der Auswertung der Fragen erstellt Ada einen personalisierten Bericht zu möglichen Ursachen für die Symptome, und deren Wahrscheinlichkeit. Ada betont dabei, keine medizinischen Diagnosen zu stellen. Die digitale Gesundheitsanwendung gibt Vorschläge für Therapiemöglichkeiten und gegebenenfalls eine Empfehlung für das Aufsuchen einer Notaufnahme.

**SkinVision:
Früherkennung
von Hautkrebs**

**Ada:
Anamnese &
Symptomanalyse**

Mindable (Mindable Health UG):

- Mindable ist eine digitale Gesundheitsanwendung zur psychotherapeutischen Behandlung von Panikstörungen und Agoraphobie. Die Inhalte von Mindable basieren auf der kognitiven Verhaltenstherapie (KVT) mit besonderem Fokus auf Konfrontationstherapie. Die KVT gilt als aktueller Goldstandard bei der Behandlung von Panikstörungen. Die digitale Gesundheitsanwendung verwendet dazu psychoedukative Videos und Texte.

Mindable:
Kognitive
Verhaltenstherapie für
Panikstörungen und
Agoraphobie

Natural Cycles (Natural Cycles Nordic AB):

- Natural Cycles ist eine digitale Gesundheitsanwendung zur Empfängnisverhütung. Die Anwendung berechnet den täglichen Fruchtbarkeitsstatus der Frau durch Eingabe der gemessenen Temperatur. Der Algorithmus von Natural Cycles analysiert die Basaltemperatur der Frau, erkennt Muster im Menstruationszyklus und ermittelt einen täglichen Fruchtbarkeitsstatus.

Natural Cycles:
Empfängnisverhütung
durch berechneten
Fruchtbarkeitsstatus

Preventicus Heartbeat (Preventicus GmbH):

- Preventicus Heartbeat ist eine digitale Gesundheitsanwendung zur Erkennung von Störungen des Herzrhythmus und Vorbeugung von Schlaganfällen. Die Anwendung nutzt dazu die Smartphonekamera, um den Herzrhythmus aufzuzeichnen. Dazu wird Licht aus dem Blitzlicht des Smartphones in die Fingerkuppe eingestrahlt. Aus den aufgezeichneten Pulskurven ermittelt der Algorithmus das Auftreten von Extrasystolen beziehungsweise das Vorhandensein einer absoluten Arrhythmie mit Verdacht auf Vorhofflimmern.

Preventicus Heartbeat:
Erkennung von Störungen
des Herzrhythmus

DiabiLive (MirambeauAppCare):

- DiabiLive ist eine digitale Gesundheitsanwendung die dem Management von insulinpflichtigem Diabetes Mellitus dient. Die Anwendung verfügt über einen Insulindosis-Rechner, dieser basiert auf dem Behandlungsprotokoll des Diabetologen. Zur Berechnung werden gemessene Blutzuckerwerte, zugeführte Kohlenhydrate und die geplante körperliche Betätigung herangezogen. Bei Überschreiten von Schwellenwerten sendet die digitale Gesundheitsanwendung automatisch SMS- und E-Mail-Benachrichtigungen gemäß ihrer Konfiguration an das Umfeld der Patient*innen.

DiabiLive:
Management von
insulinpflichtigem
Diabetes Mellitus

Kaia (Kaia Health Software GmbH):

- Kaia ist eine digitale Gesundheitsanwendung für Menschen mit chronischen Rückenschmerzen oder chronisch obstruktiver Lungenerkrankung (COPD). Kaia bietet dazu medizinische Therapieprogramme bestehend aus Wissensseinheiten, Bewegungsübungen und Entspannungstechniken. Mit einem auf künstlicher Intelligenz basierendem Bewegungscoach werden die Bewegungsausführungen der Anwender*innen analysiert und beurteilt. Dadurch soll eine korrekte Ausführung der Bewegungsübungen unterstützt werden.

Kaia:
Therapieprogramme
für chronische
Rückenschmerzen
und COPD

Moodpath (MindDoc Health GmbH):

- Moodpath ist eine digitale Gesundheitsanwendung zum Erkennen von Depressionen. Mittels digitalem Symptومتagebuch dokumentieren die Anwender*innen dreimal täglich Fragen zum Wohlbefinden. Moodpath wertet anhand der Antworten aus, ob Symptome einer Depression vorhanden sind. Diese Auswertung beinhaltet auch eine Handlungsempfehlung und einen Bericht für die Therapeut*in.

Moodpath:
Digitales
Symptومتagebuch,
Erkennen von Depression

Tinnitracks: Elektroakustische Therapie für Tinnitus	Tinnitracks (Sonormed GmbH): ■ Tinnitracks ist eine digitale Gesundheitsanwendung für Patient*innen mit Tinnitus. Die Tinnitracks Basis-Therapie besteht aus einem Lern- und Übungsprogramm zur Linderung der Belastung durch den Tinnitus. Die Tinnitracks Neuro-Therapie dient der elektroakustischen Aufbereitung von Musik (tailor-made notched music training) und dient der Behandlung von chronischem, tonalen Tinnitus.
M-sense: Digitaler Kopfschmerzkalender, Therapieprogramm für Migräne	M-sense (Newsenselab GmbH): ■ M-sense ist eine digitale Gesundheitsanwendung für Patient*innen mit Migräne. Mittels digitalem Kopfschmerztagebuch können Anwender*innen ihre Migräneattacken dokumentieren und mögliche Auslöser analysieren. Die Anwendung bietet Entspannungsverfahren wie progressive Muskelentspannung, autogenes Training und Atem-Meditation an.
mySugr: Management von insulinpflichtigem Diabetes Mellitus	mySugr (mySugr GmbH): ■ mySugr ist eine digitale Gesundheitsanwendung für Patient*innen mit Diabetes. Mit einem digitalen Diabetestagebuch können Patient*innen ihre Daten erfassen und auswerten. Die Anwendung kann mit verschiedenen Blutzuckermessgeräten gekoppelt werden, um gemessene Blutzuckerwerte automatisch an die Anwendung zu übertragen. Die Anwendung bietet einen Überblick über die gemessenen Werte, die Daten können auf Wunsch mit den Therapeut*innen geteilt werden.

Tabelle 3.2-1: Digitale Gesundheitsanwendungen – Anbieter, Funktionen und Zielgruppe (Quelle: MTRC Bericht [24], Handsuche auf Websites)

Bezeichnung	Hersteller, Land der Registrierung	Verfügbarkeit App Stores	Input, Output, Algorithmus	Anwendungsgebiet	Zielgruppe
SkinVision	SkinVision B.V. Niederlande	Apple App Store (AT): ja Google Play Store (AT): ja	Input: Bildaufnahme Output: Information & Diagnostik Algorithmus: adaptiv	Dermatologie; zur Erkennung von Hautkrebs (Melanom, Basalzellkarzinom, Plattenepithelkarzinom)	Anwender*innen (Mindestalter 18 Jahre)
Ada	Ada Health GmbH Deutschland	Apple App Store (AT): ja Google Play Store (AT): ja	Input: Texteingabe (Eigenanamnese) Output: Information & Diagnostik Algorithmus: adaptiv	Alle Fachgebiete der Medizin	Patient*innen (Mindestalter 16 Jahre)
Mindable	Mindable Health UG Deutschland	Apple App Store (AT): ja Google Play Store (AT): ja	Input: Texteingabe Output: Information & Therapie Algorithmus: keine Angabe	Psychiatrie; zur psychotherapeutischen Behandlung von Panikstörungen und Agoraphobie mittels kognitiver Verhaltenstherapie	Patient*innen (Mindestalter 16 Jahre), Therapeut*innen
Natural Cycles	Natural Cycles Nordic AB Schweden	Apple App Store (AT): ja Google Play Store (AT): ja	Input: Messwert (Temperatur) durch Texteingabe Output: Information Algorithmus: adaptiv?	Gynäkologie; zur Verhütungsmethode mittels Berechnung des Fruchtbarkeitsstatus	Anwender*innen (Mindestalter 18 Jahre)
Preventicus Heartbeat	Preventicus GmbH Deutschland	Apple App Store (AT): ja Google Play Store (AT): ja	Input: Messwert (Herzrhythmus) durch Smartphonekamera Output: Information, Diagnostik Algorithmus: adaptiv?	Innere Medizin; zur Diagnostik von Herzrhythmusstörungen und Vorbeugung von Schlaganfällen. Die Pulskurve wird mittels Smartphonekamera und Blitzlicht auf die Fingerkuppe gemessen	Anwender*innen
DiabiLive	Mirambeau AppCare Frankreich	Apple App Store (AT): ja Google Play Store (AT): ja	Input: Messwerte (Blutzucker, zugeführte Kohlenhydrate, geplante körperliche Aktivität) durch Texteingabe Output: Information, Therapie Algorithmus: adaptiv?	Innere Medizin; zum Management von Diabetes Mellitus Typ 1/2 mittels Insulindosisrechner	Patient*innen (Mindestalter 7 Jahre), Therapeut*innen
Kaia	Kaia Health Software GmbH Deutschland	Apple App Store (AT): ja Google Play Store (AT): ja	Input: Bildaufnahme (Haltung, Bewegungsmuster) mittels Smartphonekamera Output: Information, Therapie Algorithmus: adaptiv?	Orthopädie/Innere Medizin, Bewegungs- und Entspannungseinheiten für Patient*innen mit chronischen Rückenschmerzen oder chronisch obstruktiver Lungenerkrankung (COPD)	Patient*innen (Mindestalter 16 Jahre)
Moodpath	MindDoc Health GmbH Deutschland	Apple App Store (AT): ja Google Play Store (AT): ja	Input: Texteingabe Output: Diagnostik Algorithmus: fixiert?	Psychiatrie; zum Erkennen von Depressionen mittels täglichem Fragebogen zum Wohlbefinden	Anwender*innen, Therapeut*innen
Tinnitracks	Sonormed GmbH Deutschland	Apple App Store (AT): ja Google Play Store (AT): ja	Input: Audio (Musik) Output: Therapie Algorithmus: fixiert?	Oto-Rhino-Laryngologie (HNO); zur Behandlung von Tinnitus	Patient*innen, Therapeut*innen
M-Sense	Newsenselab GmbH Deutschland	Apple App Store (AT): ja Google Play Store (AT): ja	Input: Texteingabe Output: Information, Therapie Algorithmus: keine Angabe	Neurologie; zur Behandlung von Migräne (Dokumentation von Triggern, nicht-medikamentöse Therapiemethoden)	Patient*innen, Therapeut*innen
mySugr	mySugr GmbH (Roche Holding AG) Österreich	Apple App Store (AT): ja Google Play Store (AT): ja	Input: Messwerte (Blutzucker, zugeführte Kohlenhydrate, körperliche Aktivität) durch Texteingabe oder durch gekoppeltes Blutzuckermessgerät Output: Information, Benachrichtigungen, Therapie Algorithmus: adaptiv?	Innere Medizin; zum Management von Diabetes Mellitus Typ 1/2, mit verschiedenen Blutzuckermessgeräten koppelbar	Patient*innen, Therapeut*innen

3.2.2 Zusammenfassung und Diskussion der ausgewählten digitalen Gesundheitsanwendungen

unterschiedliche Funktionen refundierte Apps: Information + Diagnostik für Nutzer*innen	Digitale Gesundheitsanwendungen können in verschiedenen Stufen des Behandlungspfades Unterstützung bieten [10]: ■ Information, Früherkennung, Prävention, Diagnostik, Behandlungsempfehlung, Therapie, Nachsorge, Monitoring, Selbstmanagement
Information, Monitoring + Diagnostik für Patienten*innen erhobene Informationen (Input) führt zu Ergebnis (Output)	Die hier identifizierten und aufgrund von bereits vorliegenden Refundierungen ausgewählten Anwendungen nehmen neben der Funktion der reinen Informationen auch diagnostische und therapeutische Funktionen wahr. Diese digitalen Gesundheitsanwendungen verwenden dafür verschiedene Inputs (Informationen) der Anwender*innen. Diese können aktiv von den Anwender*innen an die digitale Gesundheitsanwendung übermittelt oder passiv von der digitalen Gesundheitsanwendung erhoben werden. Folgende Arten von Input sind bislang in Verwendung: ■ Messwerte/Laborwerte (manuelle Übertragung außerhalb der digitalen Gesundheitsanwendung erhobener Messwerte, oder Erhebung der Messwerte mittels externen Sensoren): Natural Cycles (Messung der Temperatur zur Berechnung des Fruchtbarkeitsstatus), Preventicus Heartbeat (Messung der Herzfrequenz und des Herzrhythmus), DiabiLive (Messung des Blutzuckers), mySugr (Messung des Blutzuckers) ■ Texteingabe, Eingaben über den Touchscreen: Ada, Mindable, Moodpath, M-Sense ■ Bildaufnahme, Aufnahme von Video: SkinVision (Bildaufnahme von Muttermalen), Kaia (Videoaufnahmen von Körperhaltung und Bewegungsmustern) ■ Tonaufnahme, Aufnahme von Sprache, Audio-Input: Tinnitracks ■ Sonstige Arten von Input (beispielsweise Standortdaten der Anwender*innen oder Schrittzähler)
Output: informativ, diagnostisch oder therapeutisch	Anhand der verwendeten Inputs bietet die digitale Gesundheitsanwendung den Anwender*innen Output der informativ, diagnostisch oder therapeutisch sein kann.
Fragen zu Haftung, Korrektheit der Information	Die in diesem Bericht identifizierten und ausgewählten digitalen Gesundheitsanwendungen unterscheiden sich in ihren Anwendungsgebieten, Zielgruppen (gesunde Nutzer*innen oder Patient*innen) und Funktionen. Diese haben Bedeutung für eine Evaluation zu Haftung bezüglich Korrektheit der Informationen, Organisation der Kommunikation mit Therapeut*innen sowie Entscheidungen für und Auswahl von therapeutischen Maßnahmen:
mit oder ohne Einbindung von Therapeut*innen	■ Informationen und Diagnostik an Nutzer*innen ohne Einbindung von Therapeut*innen: Ada, SkinVision, Natural Cycles, Kaia ■ Informationen, Monitoring und Diagnostik an Patient*innen ohne Einbindung von Therapeut*innen: M-Sense, Preventicus Heartbeat ■ Informationen, Monitoring und Diagnostik an Patient*innen mit Einbindung von Therapeut*innen: Mindable, DiabiLive, Moodpath, Tinnitracks, mySugr

In Tabelle 3.2-2 finden sich Angaben zur Zulassung (CE-Zertifizierung) sowie Informationen zu Preismodellen und Refundierungsmodellen aus ausgewählten Ländern. Die Kosten bei Nutzung für Anwender*innen digitaler Gesundheitsanwendungen lassen sich wie folgt unterscheiden (Einteilung der Verkaufsmodelle nach CHARISMHA-Bericht [9]):

- Kostenfrei (Kosten indirekt durch bestimmte Institutionen, private Entwickler*innen, Investor*innen getragen): Ada, Mindable
- Indirekt kostenpflichtig (Werbefinanziert)
- In-App-Verkauf: erweiterte Funktionen nach Bezahlung freigeschalten, zum Beispiel kostenpflichtige Zusatzkurse in der digitalen Gesundheitsanwendung Moodpath
- Kostenpflichtig (einmalige Gebühr bei Registrierung, danach uneingeschränkte Nutzung im Sinne einer „Flat rate“)
- Abonnements (beispielsweise monatliche oder jährliche Gebühr): SkinVision, NaturalCycles, Preventicus Heartbeat, DiabiLive, Kaia, Tinnitracks, M-Sense, mySugr

unterschiedliche Verkaufsmodelle: kostenfrei, indirekt oder direkt kostenpflichtig, Abonnements

Darüber hinaus existieren kostenfreie, gegebenenfalls zeitlich eingeschränkte Demoversionen. Ein Beispiel hierfür ist die digitale Gesundheitsanwendung Natural Cycles mit einer kostenlosen Demoversion, für die Vollversion fallen Abonnement Kosten an.

kostenfreie Demoversionen

Ein weiteres Verkaufsmodell ist eine Gebühr pro Nutzung im Sinne eines „pay-per-use“. Ein Beispiel hierfür ist die digitale Gesundheitsanwendung SkinVision, bei der die Anwender*innen anstelle eines Abonnements auch pro Bildauswertung bezahlen können.

Gebühr pro Nutzung

Eine im Jahr 2015 durchgeführte Auswertung von 641 zufällig gewählten Gesundheit- und Medizin-Apps zeigte, dass der Großteil der digitalen Gesundheitsanwendungen kostenfrei angeboten wird (31,9 % der iOS Apps aus dem Apple App Store, 44,3 % der Android Apps aus dem Google Play Store). Das Kostenmodell der Abonnements hingegen wurde von den Hersteller*innen zu diesem Zeitpunkt seltener gewählt (3,9 % der iOS Apps aus dem Apple App Store, 0,7 % der Android Apps aus dem Google Play Store) [9].

Großteil kostenfrei verfügbar

Die verschiedenen Refundierungsmodelle gliedern sich folgendermaßen:

- Jährliches Kontingent für die Nutzung digitaler Gesundheitsanwendung. Die Auswahl treffen die Anwender*innen: derzeit in keinem Land auf nationaler Ebene implementiert.
- Ermäßigungen für ausgewählte digitale Gesundheitsanwendungen: zum Beispiel 50 % Refundierung der Kosten von SkinVision durch die Partena Ziekenfonds Versicherung in Belgien.
- Gesamtübernahme der Kosten: zum Beispiel Gesamtkosten der digitalen Gesundheitsanwendung M-Sense durch Ottonova oder IKK Südwest in Deutschland refundiert.

Refundierung: Kontingent für Nutzung, Ermäßigung oder Gesamtübernahme

Diese Refundierungsmodelle können an eine Diagnosestellung gekoppelt sein, oder nur für bestimmte Bevölkerungsgruppen oder Risikogruppen gelten. Ein Beispiel hierfür ist die digitale Gesundheitsanwendung Preventicus Heartbeat für welche 10 % der Kosten durch die IKK Suedwest für bestimmte Altersgruppen und Diagnosen nach Antrag bei der Versicherung refundiert werden.

an Diagnose oder Bevölkerungsgruppe gekoppelt

Refundierung zeitlich eingeschränkt	Die Refundierung kann zeitlich eingeschränkt erfolgen. So refundieren beispielsweise die AOK Nordost, Barmer und PbeaKK in Deutschland die Gesamtkosten der digitalen Gesundheitsanwendung Kaia beschränkt auf 12 Monate.
DiGA in Österreich verfügbar, CE-Kennzeichnung	Die hier beschriebenen digitalen Gesundheitsanwendungen sind in den gängigen App Stores (Apple App Store und Google Play Store) in Österreich verfügbar. Sie sind alle zugelassen und verfügen dementsprechend über eine CE-Kennzeichnung. Die digitalen Gesundheitsanwendungen richten sich alle an gesunde Nutzer*innen, aber auch Patient*innen, teilweise zusätzlich an Therapeut*innen.
verschiedene medizinische Fachgebiete	Die elf in diesem Bericht identifizierten und ausgewählten digitalen Gesundheitsanwendungen fallen in verschiedene medizinische Fachgebiete: Innere Medizin, Orthopädie, Neurologie, Psychiatrie, Oto-Rhino-Laryngologie, Gynäkologie, Dermatologie und Allgemeinmedizin.
verschiedene Inputs und Outputs	Die verwendeten Inputs reichen von Text-, Bild-, Video- und Toneingaben bis zu Eingaben über den Touchscreen. Eine der beschriebenen digitalen Gesundheitsanwendungen verwendet die Blitzlichtfunktion des Smartphones. Die ausgegebenen Outputs fallen in die Kategorien Information, Diagnostik und Therapie. Einige der digitalen Gesundheitsanwendungen verwenden Benachrichtigungen zur Erinnerung oder bei Messwerten außerhalb der Norm.
Funktionsweise der Algorithmen nicht beschrieben	Die genaue Funktionsweise der verwendeten Algorithmen ist nie transparent beschrieben, dadurch ist eine eindeutige Zuordnung zu fixierten oder adaptiven Algorithmen nicht immer möglich.

Tabelle 3.2-2: Digitale Gesundheitsanwendungen – Versicherungen, Länder, Preis-/Refundierungsmodell (Quelle: MTRC Bericht [24], Handsuche auf Websites)

Bezeichnung	CE-Kennzeichnung	Preismodell	Land der Refundierung	Refundierende Versicherungen	Refundierungsmodell
SkinVision	Ja	Kostenloser Download der App. Pro Bildauswertung 6,99 €, 3-monatiges Abonnement 29,99 €, jährliches Abonnement 49,99 €	Belgien	Partena Ziekenfonds Versicherung	50 % Refundierung
			UK	Clinical Comission Groups	10£-20£ (je nach Nutzung)
			Deutschland	Central Krankenversicherung,	n/a 50 % Refundierung
			Niederlande	CZ, Just und NN Versicherung	Gesamtkosten refundiert
Ada	Ja	kostenlos	Deutschland	Fast-Track-Verfahren für digitale Gesundheitsanwendungen beantragt	Fast-Track-Verfahren für digitale Gesundheitsanwendungen beantragt
Mindable	Ja	kostenlos	Deutschland	Fast-Track-Verfahren für digitale Gesundheitsanwendungen beantragt	Fast-Track-Verfahren für digitale Gesundheitsanwendungen beantragt
Natural Cycles	Ja (seit 2017)	kostenlose Demoversion. Vollversion monatliches Abonnement 8,99 €, jährliches Abonnement 64,99 €	beantragt?	n/a	n/a
Preventicus Heartbeat	Ja	kostenlose Demoversion. Vollversion wöchentliches Abonnement 2,29 € monatliches Abonnement 5,49 €, jährliches Abonnement 27,99 €	Deutschland	IKK Suedwest	Co-pay 10 % für bestimmte Altersgruppen und Diagnosen nach Antrag bei der Versicherung
DiabiLive	Ja (seit 2019)	Monatliches Abonnement 25 €, jährliches Abonnement 250 €	Frankreich	Mutuelle PréviFrance	Kosten des Abonnements teilweise übernommen
Kaia	Ja	Monatliches Abonnement 99 €	Deutschland	1. AOK Nordost, Barmer, PbeaKK 2. HUK Coburg	1. Gesamtkosten refundiert (beschränkt auf 12 Monate) 2. Gesamtkosten refundiert (ohne zeitliche Einschränkung)
Moodpath	Ja	Kostenlose Basisversion, Zusatzkurse über die App kaufbar	Deutschland	Audi BKK, Barmer, ARAG, Hallesche, Signal Iduna, AXA, Gothaer	Gesamtkosten refundiert
Tinnitracks	Ja	Basisversion nur für einzelne Versicherungen. Vollversion (Neurotherapie) monatliches Abonnement 19 €	Deutschland	TK Hamburg, Bayern, Nordrhein-Westfalen, Hessen und Baden-Württemberg	Verwendung der Basisversion möglich
				63 gesetzliche Krankenkassen	Gesamtkosten der Neurotherapie Vollversion
M-Sense	Ja	3-monatiges Abonnement 89,99 €, 6-monatiges Abonnement 174,99 €, 12-monatiges Abonnement 329,99 €	Deutschland	Ottonova, IKK Südwest; Fast-Track-Verfahren für digitale Gesundheitsanwendungen beantragt	Gesamtkosten refundiert
mySugr	Ja	Kostenlose Basisversion. Vollversion monatliches Abonnement 2,99 €, mySugr Coach jährlich 357 €, mySugr Package jährlich 999 €	Deutschland	AOK Bayern, Barmenia, Gothaer, Union Krankenversicherung, Versicherungskammer Bayern, DKKV, AOK Nordwest	mySugr Package bis zu 100 % refundiert
				Fast-Track-Verfahren für digitale Gesundheitsanwendungen beantragt	

3.3 Evidenz der ausgewählten digitalen Gesundheitsanwendungen

Evidenz: Zeitpunkt der Publikation, Land, Studiendesign und Anzahl der Stud.-Teilnehmer*innen

kein Risk-of-Bias Assessment

Im folgenden Abschnitt wird die identifizierte Evidenz zu den elf ausgewählten digitalen Gesundheitsanwendungen beschrieben. Es wurde dazu auf den Webseiten der App-Anbieter nach publizierten Studien gesucht und um eine Handsuche in Pubmed ergänzt. Aufgrund der überschaubaren Anzahl vorhandener Studien zu den ausgewählten digitalen Gesundheitsanwendungen erfolgte keine Studienselektion, es werden alle identifizierten Studien im folgenden Abschnitt nach Zeitpunkt der Publikation, Land der Durchführung, gewähltes Studiendesign und Anzahl Studienteilnehmer*innen beschrieben. Ein „critical appraisal“ (Risk-of-Bias Assessment) wurde nicht durchgeführt, weil die Bewertung der digitalen Gesundheitsanwendungen nicht im Zentrum dieser Arbeit steht. Von einer Beschreibung der Detailergebnisse zu den einzelnen Studien wurde aufgrund potentieller Bias-Risiken Abstand genommen.

3.3.1 Verfügbare Evidenz zu den ausgewählten digitalen Gesundheitsanwendungen

Studien zu DiGA unterliegen hoher Dynamik

Die hier beschriebenen Studien stellen einen Versuch dar, das aktuelle Bild der Evidenz zu den jeweiligen digitalen Gesundheitsanwendungen zu veranschaulichen. Da neben der Entwicklung und Implementierung der digitalen Gesundheitsanwendungen auch die begleitenden Studien einer hohen Dynamik unterliegen, kann die nachfolgende Darstellung nur einen Ausschnitt und keinesfalls eine vollständige Übersicht über die derzeit vorhandene Evidenz zu den digitalen Gesundheitsanwendungen bieten.

SkinVision (SkinVision B.V.)

Tabelle 3.3-1: Publierte Studien zu SkinVision (CE-Kennzeichnung: 2014; derzeit geltende Risikoklasse nach MDD laut Hersteller: k.A.): <https://www.skinvision.com/hcp/>

Publikation/Jahr	Land	Studiendesign	n Patient*innen
Maier 2015 [49]	DE	Prospektive Fall-Kontrollstudie	195 Hautläsionen
Thissen 2017 [50]	NL	(Retrospektive) Fall-Kontrollstudie	108 Hautläsionen
Udrea 2020 [51]	NL	Re-Analyse [49] + Anwender-Datenbank	285 Hautläsionen
de Carvalho 2019 [52]	NL	Review	k.A.

k.A. = keine Angabe

SkinVision

Für die digitale Gesundheitsanwendung SkinVision konnten drei Studien zur Bewertung der Sensitivität und Spezifität des Algorithmus sowie ein Review zur Beschreibung der Entwicklung des verwendeten Algorithmus identifiziert werden:

3 Studien zu Sensitivität/Spezifität des Algorithmus

- In einer von Maier et al. im August 2015 publizierten prospektiven Fall-Kontrollstudie an 195 Hautläsionen wurde die Sensitivität und Spezifität der digitalen Gesundheitsanwendung beim Erkennen von malignen Melanomen gemessen. Verglichen wurden die Ergebnisse mit den Diagnosen durch 2 Dermatolog*innen, welche (unabhängig

voneinander) eine klinische Diagnostik derselben Hautläsionen vornahmen. Als Gold-Standard der Diagnostik wurde mit den histopathologischen Ergebnissen der Läsionen verglichen [49].

- In einer von Thissen et al. im Dezember 2017 publizierten Fall-Kontrollstudie an 108 Hautläsionen wurde die Sensitivität und Spezifität der digitalen Gesundheitsanwendung gemessen. Verglichen wurde mit den histopathologischen Ergebnissen der Läsionen [50].
- In einer von Udrea et al. im März 2020 publizierten Studie wurde an 285 Hautläsionen die Sensitivität und Spezifität des Algorithmus zum Erkennen von malignem Hautkrebs evaluiert [51].
- In einem von de Carvalho et al. im Januar 2019 publizierten Review über die Entwicklungen am Gebiet der Smartphone-basierten Erkennung von Hautkrebs wurden Angaben zur Funktionsweise des verwendeten Algorithmus gemacht, so wurde bis 2018 von SkinVision ein regel-basierter Algorithmus mit der Technik „fractal image analysis“ verwendet, ab 2018 wurde dieser Algorithmus durch einen auf machine-learning basierendem Algorithmus für die Bildverarbeitung und Klassifikation von Hautläsionen abgelöst [52].

Ada (Ada Health GmbH)

Tabelle 3.3-2: Publierte Studien zu ADA (CE-Kennzeichnung: 2016; derzeit geltende Risikoklasse nach MDD laut Hersteller: I)

Publikation/Jahr	Land	Studiendesign	n Patient*innen
Ronicke 2019 [53]	DE	Retrospektive Fall-Kontrollstudie	93 Patient*innen

Für die digitale Gesundheitsanwendung Ada konnte eine retrospektive Studie identifiziert werden:

- In einer von Ronicke et al. im März 2019 publizierten Studie wurde die Präzision in der Diagnostik seltener Erkrankungen von Ada DX (Prototyp des diagnostischen Entscheidungsunterstützungssystems der digitalen Gesundheitsanwendung Ada) untersucht. Die Studie wurde am Zentrum für seltene Erkrankungen an der Medizinischen Hochschule Hannover (Deutschland) durchgeführt. Für die Studie wurden Patient*innen mit einer bestätigten seltenen Erkrankung und vorhandenem Diagnosedatum in der elektronischen Gesundheitsakte ausgewählt. Von 113 identifizierten Patient*innen wurden 93 in die Studie eingeschlossen. Von diesen Patient*innen wurden die Symptome aus der elektronischen Gesundheitsakte für die Symptomanalyse mit Ada DX herangezogen. Anschließend wurde untersucht, wie häufig die vorliegende Diagnose der seltenen Erkrankung mit den Vorschlägen der Differentialdiagnostik des Algorithmus von Ada DX übereinstimmte [53].

Ada

retrospektive Studie zu Diagnostik seltener Erkrankungen

Mindable (Mindable Health UG)

Tabelle 3.3-3: Laufende Studien zu Mindable (CE-Kennzeichnung: vorhanden, keine Angabe zum Zeitpunkt; derzeit geltende Risikoklasse nach MDD laut Hersteller: 1) <https://www.mindable.health/en/science/>

Publikation/Jahr	Land	Studiendesign	n Patient*innen
Christoph-Dornier-Stiftung: Studie 1 [54]	DE	k.A.	k.A.
Christoph-Dornier-Stiftung: Studie 2 [54]	DE	k.A.	k.A.
Universitätsklinik Bonn [54]	DE	k.A.	k.A.

k.A. = keine Angabe

Mindable

Für die digitale Gesundheitsanwendung Mindable konnten drei aktuell laufende Studien identifiziert werden:

aktuell 3 laufende Studien

- In Zusammenarbeit mit der Christoph-Dornier-Stiftung für klinische Psychologie wird die Wirksamkeit von Mindable in zwei Studien untersucht. In der ersten Studie sollen die Effekte der Nutzung der digitalen Gesundheitsanwendung während der Wartezeit auf einen Psychotherapieplatz und deren Einfluss auf die Behandlungseffizienz dieser anschließenden Psychotherapie untersucht werden. In der zweiten Studie sollen die Effekte der (zur Psychotherapie) begleitenden Nutzung der digitalen Gesundheitsanwendung auf die Symptome untersucht werden. Zum genauen geplanten Studiendesign dieser beiden Studien lagen noch keine Angaben vor [54].
- In Zusammenarbeit mit der Universitätsklinik Bonn ist eine Studie zur Nutzung von Mindable bei stationären psychiatrischen Patient*innen mit Komorbiditäten geplant. Zum genauen geplanten Studiendesign lagen noch keine Angaben vor [54].

Natural Cycles (Natural Cycles Nordic AB)

Tabelle 3.3-4: Publierte Studien zu Natural Cycles (CE-Kennzeichnung: 2017; derzeit geltende Risikoklasse nach MDD laut Hersteller: 2b) <https://www.naturalcycles.com/research>, <https://help.naturalcycles.com/hc/en-us/articles/360003368593-Research-at-Natural-Cycles>

Publikation/Jahr	Land	Studiendesign	n Anwender*innen
Berglund 2016 [55]	SE	Retrospektive Beobachtungsstudie	4.054 Frauen
Berglund 2017 [56]	SE	Prospektive Beobachtungsstudie	22.875 Frauen
Bull 2019 [57]	SE	Prospektive Beobachtungsstudie	16.331 Frauen
Berglund 2019 [58]	SE	Prospektive Beobachtungsstudie	2.874 Frauen
Bull 2019 [59]	SE	Retrospektive Beobachtungsstudie	124.648 Frauen
Kleinschmidt 2019 [60]	SE	Prospektive Beobachtungsstudie	42.579 Frauen

Natural Cycles ist in Europa als Verhütungsmittel in der Risikoklasse 2b zugelassen. Für die digitale Gesundheitsanwendung Natural Cycles konnten sechs Studien identifiziert werden:

- In einer von Berglund et al. im März 2016 publizierten retrospektiven Beobachtungsstudie wurde die Wirksamkeit von Natural Cycles in Bezug auf Empfängnisverhütung untersucht. Zur Messung der Empfängnisverhütung wurde der Pearl-Index (Maß zur Beschreibung des Anteils an Frauen die trotz Verwendung einer bestimmten Verhütungsmethode innerhalb eines Jahres schwanger werden) von 4.054 Anwenderinnen, entsprechend 2.085 Frauenjahren, berechnet. Die Studie wurde ohne vergleichende Kontrollgruppe durchgeführt [55].
- In einer von Berglund et al. im Dezember 2017 publizierten prospektiven Beobachtungsstudie wurde die Wirksamkeit von Natural Cycles in Bezug auf Empfängnisverhütung untersucht. Dabei wurden die Messungen von 22.875 Anwenderinnen, entsprechend 18.548 Frauenjahren, nach „typical use“ Pearl-Index und „perfect use“ Pearl-Index untersucht. Die Studie wurde ohne vergleichende Kontrollgruppe durchgeführt [56].
- In einer von Bull et al. im März 2019 publizierten prospektiven Beobachtungsstudie wurde das Risiko einer ungewollten Schwangerschaft von 16.331 Anwenderinnen, welche Natural Cycles verwendeten, untersucht [57].
- In einer von Berglund et al. im Juni 2019 publizierten prospektiven Beobachtungsstudie wurde die Wirksamkeit von Natural Cycles in Bezug auf die Fekundabilität (die Wahrscheinlichkeit, eine Schwangerschaft pro Menstruationszyklus zu erreichen) untersucht. An der Studie haben 2874 Frauen mit dem Wunsch einer Schwangerschaft teilgenommen. Die Studienteilnehmerinnen wurden je nach vorangegangener Empfängnisverhütung in zwei Gruppen geteilt (vorangegangene Empfängnisverhütung mittels Natural Cycles oder vorangegangene Empfängnisverhütung mittels hormoneller Kontrazeption). Die durchschnittliche Zeit zur Schwangerschaft (time to pregnancy, TTP) wurde berechnet [58].
- In einer von Bull et al. im August 2019 publizierten retrospektiven Beobachtungsstudie wurden 612.613 Menstruationszyklen von 124.648 Anwenderinnen der Natural Cycles Anwendung analysiert um Zusammenhänge zwischen Zykluslänge, Alter und Body-Mass-Index (BMI) zu identifizieren [59].
- In einer von Kleinschmidt et al. im November 2019 publizierten prospektiven Beobachtungsstudie wurde die Identifikation der fruchtbaren Tage durch Natural Cycles verglichen mit der „rhythm method“ (Kalendermethode nach Knaus-Ogino) und der „standard days method“ (vereinfachte Kalendermethode). Dabei wurden 26.626 Menstruationszyklen analysiert und richtig-positive („green days“) von falsch-positiven („wrong green days“) Parametern unterschieden [60].

Natural Cycles

prospektive & retrospektive Beobachtungsstudien zu Pearl-Index und Empfängniswahrscheinlichkeit

Preventicus Heartbeat (Preventicus GmbH)

Tabelle 3.3-5: Publierte Studien zu Preventicus Heartbeat (CE-Kennzeichnung: vorhanden, keine Angabe zum Zeitpunkt; derzeit geltende Risikoklasse nach MDD laut Hersteller: 2a) <https://www.preventicus.com/en/about-preventicus/studien/>

Publikation/Jahr	Land	Studiendesign	n Anwender*innen
Koenig 2016 [61]	DE/CH	Prospektive Fall-Kontroll-Studie	68 Anwender*innen
Krivoshei 2017 [62]	CH/DE	Prospektive Fall-Kontroll-Studie	80 Patient*innen
Weber 2017 [63] (nur Poster)	DE/CH	Prospektive Fall-Kontroll-Studie	46 Patient*innen
Karim 2017 [64] (nur Poster)	UK	Prospektive Fall-Kontroll-Studie	140 Patient*innen
Brasier 2019 [65]	CH/DE	Prospektive Fall-Kontroll-Studie	672 Patient*innen
Dörr 2019 [66]	CH/DE	Prospektive Fall-Kontroll-Studie	672 Patient*innen
Birkemeyer 2020 [67]	DE	Kosten-Effektivitäts-Analyse	k.A.

k.A. = keine Angabe

Preventicus Heartbeat

6 Studien zu Sensitivität/
Spezifität des Algorithmus,
1 Studie zu Kosten-
Effektivitäts-Analyse

Für die digitale Gesundheitsanwendung Preventicus Heartbeat konnten sieben Studien identifiziert werden:

- Eine im August 2016 publizierte prospektive Fall-Kontroll-Studie von Koenig et al. untersuchte die Möglichkeit die Herzfrequenz und Herzfrequenzvariabilität mittels Smartphonekamera korrekt zu messen. Dazu wurde bei 68 Studienteilnehmer*innen gleichzeitig eine EKG-Aufzeichnung und Aufzeichnung mittels Smartphonekamera durchgeführt [61].
- Eine im Mai 2017 publizierte prospektive Fall-Kontroll-Studie von Krivoshei et al. untersuchte die Erkennung von Vorhofflimmern durch den Algorithmus von Preventicus Heartbeat. Dazu wurden 80 Patient*innen für die Studie rekrutiert (40 mit diagnostiziertem Vorhofflimmern, 40 mit Sinusrhythmus). Dabei wurden verschiedene Auswertungsmodelle in der Diagnostik über die berechnete Sensitivität und Spezifität miteinander verglichen [62].
- Eine im Juni 2017 publizierte prospektive Fall-Kontroll-Studie von Weber et al. untersuchte, ob die Erkennung von Extrasystolen durch die Smartphonekamera möglich ist. Dazu wurde bei 46 Patient*innen der Puls mittels EKG und Smartphonekamera aufgezeichnet und miteinander verglichen. Die Sensitivität und Spezifität der gemessenen Extrasystolen wurde berechnet [63].
- Eine im Juni 2017 publizierte prospektive Fall-Kontroll-Studie von Karim et al. verglich die Diagnosegenauigkeit von zwei verschiedenen digitalen Gesundheitsanwendungen (BeatScanner und Preventicus Heartbeat). Während die BeatScanner App zur Diagnostik den Beschleunigungsmesser und das Gyroskop des Smartphones nutzt, verwendet die Preventicus Heartbeat App zur Diagnostik die Smartphonekamera und die Blitzlichtfunktion. Für die Studie wurden 70 Patient*innen mit Vorhofflimmern und 70 Anwender*innen mit Sinusrhythmus rekrutiert. Für beide digitale Gesundheitsanwendungen wurde die Sensitivität und Spezifität des Screenings für Vorhofflimmern berechnet [64].

- Eine im Januar 2019 publizierte prospektive Fall-Kontroll-Studie von Brasier et al. verglich den von Preventicus Heartbeat verwendeten Algorithmus mit der Diagnose durch Elektrokardiogramm (durch Kardiolog*innen ausgewertet), um zwischen Vorhofflimmern und Sinusrhythmus zu unterscheiden [65].
- Eine im Februar 2019 publizierte prospektive Fall-Kontroll-Studie von Dörr et al. verglich die Diagnosegenauigkeit von Vorhofflimmern des Preventicus Heartbeat Algorithmus (basierend auf photoplethysmographischen Signalen) mit der Diagnosegenauigkeit von Elektrokardiogrammen (durch Kardiolog*innen ausgewertet) [66].
- Eine im Juni 2020 publizierte Studie von Birkemeyer et al. erstellte eine Kosten-Effektivitäts-Analyse zum Screening von Vorhofflimmern mittels der digitalen Gesundheitsanwendung Preventicus Heartbeat aus der Perspektive deutscher gesetzlicher Krankenversicherungen [67].

DiabiLive (MirambeauAppCare)

Tabelle 3.3-6: *Publizierte Studien zu DiabiLive (CE-Kennzeichnung: 2019; derzeit geltende Risikoklasse nach MDD laut Hersteller: 2b)*
<https://www.diabilive.com/story>

Publikation/Jahr	Land	Studiendesign	n Anwender*innen
k.A.	k.A.	k.A.	k.A.

Für die digitale Gesundheitsanwendung DiabiLive konnten keine abgeschlossenen Studien identifiziert werden. Ebenso konnten keine aktuell laufende oder sich in Planung befindende Studien identifiziert werden.

**DiabiLive
ohne Evidenz**

Kaia (Kaia Health Software GmbH)

Tabelle 3.3-7: *Publizierte und aktuell laufende Studien zu Kaia (CE-Kennzeichnung: vorhanden, keine Angabe zum Zeitpunkt; derzeit geltende Risikoklasse nach MDD laut Hersteller: I)*
<https://www.kaiahealth.com/de/rechtliches/agb/>

Publikation/Jahr	Land	Studiendesign	n Anwender*innen
Huber 2017 [68]	DE	Retrospektive Beobachtungsstudie	180 Anwender*innen
Clement 2018 [69]	DE	Retrospektive Kohortenstudie	196 Anwender*innen (v0), 1.055 Anwender*innen (v1)
Toelle 2019 [70]	DE	Randomisierte kontrollierte Studie DRKS00016329	101 Anwender*innen
Brigham and Women's Hospital, Boston/MA	US	Randomisierte kontrollierte Studie NCT04290078	184 Anwender*innen
Schön Klinik Berchtesgadener Land	DE	Prospektive Kohortenstudie NCT04457843	130 Anwender*innen

v0 = Version 0.x, v1 = Version 1.x

Kaia 2 retrospektive Kohortenstudien, 1 RCT Aktuell 2 laufende Studien (prospektive Kohortenstudie, RCT)	Für die digitale Gesundheitsanwendung Kaia konnten fünf Studien identifiziert werden. Zu den abgeschlossenen Studien zählen eine randomisierte kontrollierte Studie, eine retrospektive Kohortenstudie und eine retrospektive Beobachtungsstudie. Ebenso wurden zwei aktuell laufende Studien identifiziert: ■ Eine im Dezember 2017 publizierte retrospektive Beobachtungsstudie von Huber et al. analysierte retrospektiv die anonymisierten Daten von 180 Nutzer*innen der digitalen Gesundheitsanwendung Kaia. Dabei wurde die Reduktion der Schmerzintensität nach numerischer Rating-Skala gemessen. Die Studie wurde ohne vergleichende Kontrollgruppe durchgeführt, die Autor*innen empfehlen zur Bias-Reduktion die Durchführung einer prospektiven, randomisierten und kontrollierten Studie [68]. ■ Eine im Juni 2018 publizierte retrospektive Kohortenstudie von Clement et al. untersuchte Unterschiede verschiedener App-Versionen in Bezug auf die Häufigkeit der Nutzung der digitalen Gesundheitsanwendung und verbesserte klinische Resultate. Dazu wurden frühere Versionen der digitalen Gesundheitsanwendung (0.x) mit der neueren Version der digitalen Gesundheitsanwendung (1.x) verglichen. 196 Nutzer*innen der früheren Version wurden mit den Daten von 1.055 Nutzer*innen der neueren Version verglichen. Verglichen wurde die Dauer der Nutzung der digitalen Gesundheitsanwendung, die Anzahl abgeschlossener Übungen sowie von den Anwender*innen berichtete Schmerzintensität [69]. ■ Eine im Mai 2019 publizierte randomisiert kontrollierte Studie von Toelle et al. untersuchte 101 Patient*innen mit unspezifischen Schmerzen im unteren Rückenbereich. Die Teilnehmer*innen wurden zufällig in eine Interventionsgruppe und eine Kontrollgruppe eingeteilt. Die Teilnehmer*innen der Interventionsgruppe wurden 3 Monate mit der digitalen Gesundheitsanwendung Kaia behandelt. Die Teilnehmer*innen der Kontrollgruppe erhielten sechs Einheiten Physiotherapie kombiniert mit Online-Kursen. Als primärer Endpunkt wurde die Schmerzintensität mittels numerischer Rating-Skala gemessen [70]. ■ Eine aktuell laufende Studie am Brigham and Women's Hospital (in Boston, USA) an Patient*innen mit Rückenschmerzen untersucht den Einfluss der digitalen Gesundheitsanwendung Kaia auf die Entwicklung der Schmerzintensität, körperlichen Funktionsfähigkeit und Lebensqualität nach PROMIS-10. Die Studie sieht 184 Teilnehmer*innen einer randomisiert kontrollierten Studie vor. Clinical Trials Identifier: NCT04290078. ■ Eine aktuell laufende Studie an der Schön Klinik (in Berchtesgader Land, Deutschland) an Patient*innen mit COPD, welche die digitale Gesundheitsanwendung Kaia nutzen. Ziel der Studie ist die Erfassung der Korrelation von konventionellen Fragebögen zu digitalen Messungen über die digitale Gesundheitsanwendung in Bezug auf die Krankheitsaktivität. Die Studie sieht 130 Teilnehmer*innen einer prospektiven Kohortenstudie vor. Clinical Trials Identifier: NCT04457843.
--	--

Moodpath (MindDoc Health GmbH)

Tabelle 3.3-8: Publierte Studien zu Moodpath (CE-Kennzeichnung: 2016; derzeit geltende Risikoklasse nach MDD laut Hersteller: 1) <https://mymoodpath.com/de/wissenschaft/>, <https://mymoodpath.com/en/ce/>

Publikation/Jahr	Land	Studiendesign	n Anwender*innen
Scherr 2019 [71]	k.A.	Retrospektive Beobachtungsstudie	6.675 Anwender*innen
Martinengo 2019 [72]	mn	Review von 69 Apps	k.A.

k.A. = keine Angabe, mn=multinational

Für die digitale Gesundheitsanwendung Moodpath konnten zwei publizierte Studien identifiziert werden:

- Eine im April 2019 publizierte retrospektive Beobachtungsstudie von Scherr et al. setzte sich zum Ziel zu untersuchen, ob Moodpath eine geeignete Quelle für Informationen zu psychischer Gesundheit darstellen kann. Mittels „ecological momentary assessment“ (EMA) von 6.675 Anwender*innen der digitalen Gesundheitsanwendung wurde untersucht, ob Anwender*innen mit depressiven Symptomen in der Selbstbeobachtung vermehrt nach Informationen zu psychischer Gesundheit gesucht hatten. Die Studie wurde ohne vergleichende Kontrollgruppe durchgeführt [71].
- Eine im Dezember 2019 publizierte Studie von Martinengo et al. zum systematischen Assessment zur Einhaltung klinischer Leitlinien von digitalen Gesundheitsanwendungen aus dem Bereich der psychischen Gesundheit untersuchte, wieviele der bewerteten digitalen Gesundheitsanwendungen allen klinischen Leitlinien zur Suizidprävention entsprechen. Die sechs zum Assessment herangezogenen klinischen Leitlinien zur Suizidprävention waren Monitoring von Stimmungs- und Suizidgedanken, Entwicklung von Sicherheitsplänen, Empfehlung von Aktivitäten zur Ablenkung von Suizidgedanken, Information und Aufklärung, Zugang zu unterstützenden Netzwerken und Zugang zu Notfallberatung. Moodpath war eine der 69 untersuchten digitalen Gesundheitsanwendungen dieses systematischen Assessments [72].

Moodpath

2 Publikationen:
 „Ecological momentary assessment“;
 Bewertung von Moodpath als Teil eines systematischen Assessments

Tinnitracks (Sonormed GmbH)

Tabelle 3.3-9: Publierte Studien zu Tinnitracks/TMNMT (CE-Kennzeichnung: vorhanden, keine Angabe zum Zeitpunkt; derzeit geltende Risikoklasse nach MDD laut Hersteller: 1) <http://www.tinnitracks.com/de/aktuelles/tmnmt-anbieter>, <https://www.tinnitracks.com/de/therapiewirkung>

Publikation/Jahr	Land	Studiendesign	n Anwender*innen
Okamoto 2010 [73]	DE	Randomisiert kontrollierte Studie	16 Patient*innen
Stein 2016 [74]	DE	Randomisiert kontrollierte Studie	100 Patient*innen
Löhler 2019 [75]	DE	Anwenderumfrage	117 HNO-Ärzt*innen

Für die digitale Gesundheitsanwendung Tinnitracks wurden allgemeine, nicht für die digitale Gesundheitsanwendung spezifische Studien zum verwendeten Therapieansatz mit frequenzgefilterter Musik gefunden. Spezifisch für die digitale Gesundheitsanwendung Tinnitracks wurde eine Anwenderumfrage gefunden:

Tinnitracks

Studien zum verwendeten Therapieansatz „TMNMT“, Anwenderumfrage zu Tinnitracks

- Der Hersteller verweist auf allgemeine Studien zur frequenzgefilterten Musik („tailor made notched music training“, TMNMT). 13 allgemeine, nicht Tinnitracks-spezifische Studien zu TMNMT werden auf der Website des Herstellers zitiert, darunter die Studie von Okamoto [73] und Stein [74]:
- In einer 2010 publizierten randomisiert-kontrollierten Studie von Okamoto et al. wurden 16 Patient*innen mit Tinnitus pseudorandomisiert in eine TMNMT-Gruppe und eine Placebo-Gruppe eingeteilt. Nach 12 Monaten wurde die empfundene Tinnituslautstärke beider Gruppen verglichen [73].
- In einer 2016 publizierten randomisiert-kontrollierten Studie von Stein et al. an 100 Studienteilnehmer*innen wurde der Effekt von TMNMT auf die empfundene Tinnituslautstärke (und weitere Messgrößen des Tinnitus) gemessen [74].
- Eine im November 2019 publizierte Studie von Löhler et al. befragte 457 HNO-Ärzt*innen, welche Tinnitracks zur Therapie ihrer Patient*innen einsetzten. 117 der befragten Ärzt*innen äußerten sich zu ihrem Eindruck zu Wirksamkeit und Verträglichkeit der digitalen Gesundheitsanwendung [75].

M-sense (Newsenselab GmbH)

Tabelle 3.3-10: Aktuell laufende Studien zu M-sense (CE-Kennzeichnung: vorhanden, keine Angabe zum Zeitpunkt; derzeit geltende Risikoklasse nach MDD laut Hersteller: I)
<https://www.m-sense.de/faq>, <https://www.m-sense.de/en/impressum>, <https://www.m-sense.de/forschung>

Publikation/Jahr	Land	Studiendesign	n Anwender*innen
Klinik für Neurologie, Charité	DE	Randomisierte kontrollierte Studie	k.A.
Barmer, Deutsche Telekom AG, Deutsche Post AG	DE	k.A.	k.A.

k.A. = keine Angabe

M-sense	Für die digitale Gesundheitsanwendung M-sense konnten zwei aktuell laufende Studien identifiziert werden:
aktuell 2 laufende Studien (RCT und Studiendesign o.A.)	■ Eine aktuell laufende Studie an Patient*innen mit Migräne an der Universitätsklinik für Neurologie an der Charité untersucht im Rahmen des „SMARTGEM“-Projekts Smartphone-gestützte Migränetherapie. Die Effektivität einer neuen Versorgungsform (digitale Gesundheitsanwendung kombiniert mit telemedizinischen Angeboten) soll mittels einer randomisiert kontrollierten Studie evaluiert werden. Zur geplanten Anzahl der Studienteilnehmer*innen lagen noch keine Angaben vor. ■ Eine aktuell laufende Studie an Patient*innen mit Migräne soll evaluieren, ob mit der Nutzung der digitalen Gesundheitsanwendung M-sense die Kopfschmerzstage reduziert und die Lebensqualität verbessert werden kann. Die Studie erfolgt in Kooperation mit der Barmer-Krankenkasse, der deutschen Telekom AG sowie der deutschen Post AG. Zum genauen geplanten Studiendesign lagen noch keine Angaben vor.

mySugr (mySugr GmbH)

Tabelle 3.3-11: *Publizierte Studien zu mySugr (CE-Kennzeichnung: vorhanden (2013?); derzeit geltende Risikoklasse nach MDD laut Hersteller: 1, Bolusrechner als Modul Klasse 2b)*
https://www.mysugr.com/static/assets/downloads/press_releases/201809-mysugr-BARMER.pdf,
<https://support.mysugr.com/hc/de/articles/203892086-Sind-meine-Daten-sicher>

Publikation/Jahr	Land	Studiendesign	n Anwender*innen
Miller 2014 [76]	n/a	Review	k.A.
Peng 2016 [77]	USA	Befragung, Fokusgruppe	18 Anwender*innen
Martinez 2017 [78]	USA	Review zu 85 Diabetes Apps	k.A.
Kebede 2019 [79]	DE	Befragung via Facebook	1.052 Anwender*innen (DM Typ 1), 630 Anwender*innen (DM Typ 2)
Fu 2019 [80]	USA	Randomisierte Cross-Over Studie	92 Anwender*innen
Dehong 2019 [81]	AT	Narrativer Review von 4 retrospektiven Beobachtungsstudien + Befragung	2.104 Anwender*innen, 440 Anwender*innen, 4.000 Anwender*innen, 52 Anwender*innen, 105 Anwender*innen
Quevedo 2019 [82]	ES	Systematisches Review zu 42 Apps (Qualität), 34 Apps (Benutzerfreundlichkeit)	k.A.
Mayer 2019 [83] (nur Poster)	AT	Retrospektive Beobachtungsstudie	61 Anwender*innen
Fu 2020 [84]	USA	Review zu 4 Apps	k.A.
Hospital de la Santa Creu i Sant Pau, Barcelona	ES	Randomisierte kontrollierte Studie NCT03819335	28 Anwender*innen
Steno Diabetes Center Copenhagen	DK	Randomisierte kontrollierte Studie NCT03682237	184 Anwender*innen

k.A. = keine Angabe

Für die digitale Gesundheitsanwendung mySugr konnten 3 Reviews, 3 Befragungen, 1 randomisierte Cross-Over Studie, 5 retrospektive Beobachtungsstudien sowie 2 geplante randomisiert kontrollierte Studien identifiziert werden:

- Eine im Juli 2014 publizierte Studie von Miller et al. untersuchte von verschiedenen digitalen Gesundheitsanwendungen, darunter mySugr, die verwendeten Mechanismen zur Gamifizierung (die Anwendung spieltypischer Elemente in einem spielfremden Kontext). Die Autor*innen beschreiben eine potentielle Verbesserung des Selbstmanagements chronischer Krankheiten durch Gamifizierung [76].
- Eine im September 2016 publizierte Studie von Peng et al. befragte Patient*innen mit Diabetes Mellitus Typ 2 zu ihrer Einschätzung nach Barrieren, Nutzen und Funktionen digitaler Gesundheitsanwendungen, darunter mySugr [77].
- Eine im Juli 2017 publizierte Studie von Martinez et al. untersuchte 85 digitale Gesundheitsanwendungen für Diabetes, darunter mySugr, auf deren angebotene Funktionen und Grad der Interaktivität sowie App Store/Play Store Bewertungen der Anwender*innen sowie deren vorgeschlagene Funktionen [78].
- Eine im März 2019 publizierte Befragung von Kebede et al. an 1.052 Patient*innen mit Diabetes Mellitus Typ 1 sowie 630 Patient*innen mit Diabetes Mellitus Typ 2 untersuchte den Grad der Selbstversorgung bei Nutzung digitaler Gesundheitsanwendungen [79].

mySugr

**3 Reviews,
3 Befragungen,
1 randomisierte
Cross-Over-Studie,
5 retrospektive
Beobachtungsstudien,
2 geplante RCTs**

- Eine im April 2019 publizierte randomisierte Cross-Over Studie von Fu et al. untersuchte den Einfluss von Patient*innenmerkmalen auf die Benutzerfreundlichkeit von zwei digitalen Gesundheitsanwendungen (OnTrack und mySugr) [80].
- Eine im Mai 2019 publizierte Studie von Debong et al. untersuchte Veränderungen in der glykämischen Kontrolle und Patient*innenzufriedenheit bei Nutzung von mySugr. Die Autor*innen verweisen auf retrospektive Beobachtungsstudien [81].
- Ein im Mai 2019 publiziertes systematisches Review von Quevedo et al. zu digitalen Gesundheitsanwendungen evaluierte 42 Apps nach Qualität und 34 Apps nach Benutzerfreundlichkeit. Die digitale Gesundheitsanwendung mySugr war eine der evaluierten Apps. Laut den Autor*innen geben die wenigsten digitalen Gesundheitsanwendungen wissenschaftliche Referenzen zum Inhalt, außerdem greifen die meisten digitalen Gesundheitsanwendungen auf persönliche Daten und Gerätedaten zu [82].
- Eine im Juni 2019 publizierte retrospektive Beobachtungsstudie von Mayer et al. untersuchte Veränderungen der Blutzuckerkontrolle bei Verwendung von mySugr [83].
- Eine im Juni 2020 publizierte Studie von Fu et al. untersuchte die Benutzerfreundlichkeit von 4 digitalen Gesundheitsanwendungen, darunter mySugr. Die Autor*innen fordern eine Erhöhung der Benutzerfreundlichkeit durch Tutorials, eine geeignete Suchfunktion und ästhetisches Design [84].
- Eine aktuell laufende Studie am Hospital de la Santa Creu i Sant Pau in Barcelona untersucht mittels Diabetes Empowerment Scale-Short Form (DES-SF) den Effekt von mySugr auf das Management von Diabetes. Die Studie sieht 28 Anwender*innen einer randomisiert kontrollierten Studie vor. Clinical Trials Identifier: NCT03819335.
- Eine aktuell laufende Studie am Steno Diabetes Center in Copenhagen untersucht die Zeit in Normoglykämie bei Verwendung von mySugr. Die Studie sieht 184 Anwender*innen einer randomisiert kontrollierten Studie vor. Clinical Trials Identifier: NCT03682237.

3.3.2 Zusammenfassung und Diskussion der vorliegenden Evidenz zu ausgewählten digitalen Gesundheitsanwendungen

Unterschiede bei Wahl der Studiendesigns

reine Anwenderdaten Analysen

Fall-Kontrollstudien

RCTs in rezenten Studien häufiger

In einer Zusammenschau der Evidenz der digitalen Gesundheitsanwendungen zeigt sich, dass es große Unterschiede bei der Wahl der Studiendesigns – selbst bei gleichen Risikoklassen – gibt. Neben reinen (retrospektiven oder prospektiven) Analysen der Real World Daten (RWE) der Anwender*innen (NaturalCycles, Moodpath) kamen auch (retrospektive oder prospektive) Fall-Kontrollstudien (Preventicus Heartbeat, ADA, SkinVision) zum Einsatz. Bei RWE-Datenbankanalysen werden Anwender*innen-Daten vom Anbieter erhoben und im Sinn einer Beobachtungsstudie analysiert, ohne jedoch Kenntnisse zu Charakteristika der Anwender*innen, Umstände der Anwendung, etc. zu kennen.

Randomisiert kontrollierte Studien stellen den Goldstandard zum Nachweis eines kausalen Effekts dar – dieses Studiendesign ist bei den ausgewählten digitalen Gesundheitsanwendungen nicht immer, aber in rezenten Studien (z. B. laufende Studien) häufiger gegeben.

Tabelle 3.3-12: Digitale Gesundheitsanwendung, Risikoklasse und höchste Evidenzstufe

GesundheitsApp	Zulassungsjahr	Derzeitige Risikoklasse (nach MDD, laut Hersteller)	höchste Evidenzstufe + aktuell laufende Studien
DiabiLive	2019	2b	Keine Studien
mySugr	k.A.	2b	1 RCT Crossover + 2 RCTs laufend
NaturalCycles	2017	2b	4 prospektive Beobachtungsstudien (RWE-Datenbankanalysen), 2 retrospektive Beobachtungsstudien (RWE-Datenbankanalysen)
Preventicus Heartbeat	k.A.	2a	5 prospektive Fall-Kontrollstudien
ADA	2016	1	1 retrospektive Fall-Kontrollstudie
Mindable	k.A.	1	Studien laufend, Studiendesign k.A.
Kaia	k.A.	1	1 RCT + 1 Kohortenstudie laufend + 1 RCT laufend
Moodpath	k.A.	1	1 retrospektive Beobachtungsstudie (RWE-Datenbankanalyse)
Tinnitracks	k.A.	1	2 RCTs
M-sense	k.A.	1	1 RCT laufend
SkinVision	2014	k.A.	1 prospektive Fall-Kontrollstudie, 1 retrospektive Fall-Kontrollstudie

k.A. = keine Angabe, RCT = randomisierte klinische Studie, RWE = real world evidence

Einige der hier beschriebenen digitalen Gesundheitsanwendungen verweisen auf Studien, die nicht spezifisch für die jeweilige digitale Gesundheitsanwendung durchgeführt wurden. Die Anbieter verweisen auf den Websites auf:

- allgemeine Studien zur Wirksamkeit von *Smartphone-Apps*. So verweist beispielsweise Mindable (da für die Wirksamkeit von Mindable im Speziellen noch keine Studienergebnisse vorliegen) auf die allgemeine Wirksamkeit von Smartphone-Apps in der Behandlung psychischer Erkrankung auf Studien und systematische Literaturreviews.
- allgemeine Studien zur Wirksamkeit des verwendeten *Therapieansatzes*. So verweist beispielsweise Tinnitracks auf Studien über die Wirksamkeit von Therapie mit frequenzgefilterter Musik, jedoch nicht auf Studien über Tinnitracks im Speziellen.
- Reviews, in denen gleich mehrere digitale Gesundheitsanwendungen miteinander verglichen und bewertet wurden: so war beispielsweise Moodpath eine der bewerteten digitalen Gesundheitsanwendungen eines systematischen Vergleichs.

Ergebnisse dieser Studien zur allgemeinen Wirksamkeit von digitalen Gesundheitsanwendungen bzw. Studien zur Wirksamkeit des verwendeten Therapieansatzes können in die Bewertung einer digitalen Gesundheitsanwendung miteinfließen und vor allem zur Hypothesenbildung beitragen, welche Effekte möglicherweise zu erwarten sind (und welche Endpunkte valide Aussagen über diese Effekte machen können). Alleinstehend sind sie jedoch nur bedingt geeignet um eine Aussage zur Wirksamkeit einer digitalen Gesundheitsanwendung im Speziellen zu treffen.

Eine der hier beschriebenen digitalen Gesundheitsanwendungen (DiabiLive, Risikoklasse 2b) verfügt bislang über keine Studie (weder abgeschlossen, noch in Planung) und somit keine Evidenz zum Nutzen der digitalen Gesundheitsanwendung.

**Verweis der Hersteller
auf allgemeine Studien
zu DiGA**

**Studien zur allgemeinen
Wirksamkeit von DiGA**

**DiGA ohne vorliegende
Evidenz**

Studien zum Teil mit schlechter methodischer Qualität

Ein im Februar 2020 publizierter systematischer Review zur diagnostischen Genauigkeit von digitalen Gesundheitsanwendungen zur Diagnostik von Hautkrebs identifizierte 235 dermatologische Apps, die zwischen 2014 und 2017 angeboten wurden (einige auch wieder vom Markt genommen). Im Review konnten aber nur neun Studien (u. a. zu zwei in Europa zugelassener Apps: SkinScan and SkinVision) analysiert werden: in sechs Studien wurde die diagnostische Genauigkeit mit Histologie oder Follow-up verglichen (n = 725 Hautläsionen), in drei Studien wurde mit der Einschätzung von Dermatolog*innen verglichen (n = 407 Hautläsionen). Die Autor*innen stellen den Studien insgesamt schlechte methodische Qualität aus (selektive Rekrutierung, selektive Auswahl zu evaluierender Hautläsionen) und beschreiben für SkinVision niedrige Übereinstimmung mit der Einschätzung von Dermatolog*innen [85].

Notwendigkeit für Studien für Nachfolgeprodukte nicht eindeutig

Abschließend stellt sich die Frage (wie auch bei anderen Medizinprodukten: first-in-class vs. Nachfolgeprodukte), ob für die Evidenz zur Wirksamkeit einer digitalen Gesundheitsanwendung in einem Einsatzgebiet allgemeine Studien zur Wirksamkeit des verwendeten Therapieansatzes oder einer alternativen Smartphone-Apps bei einer speziellen Indikation ausreichen, oder ob es Wirksamkeitsnachweis über die jeweilige digitale Gesundheitsanwendung im Speziellen bedarf.

3.4 Anwendung der Assessment Frameworks

Einteilung der DiGA in Risikoklassen und Evidenzstufen

Im folgenden Abschnitt werden die in diesem Projekt ausgewählten Assessment Frameworks auf die ausgewählten digitalen Gesundheitsanwendungen angewendet. Zunächst folgt eine Beschreibung mit allgemeinen Erläuterungen zur Anwendbarkeit der einzelnen Frameworks. Im darauffolgenden Schritt erfolgt eine Zuteilung der ausgewählten digitalen Gesundheitsanwendungen in die derzeitige Risikoklasse nach Hersteller sowie die jeweilige voraussichtliche Risikoklasse nach 2017/745 MDR und Einteilung nach Evidenzstufen von NICE.

Besonderheiten in der Evaluation von DiGA

Anschließend werden die Assessment Frameworks, soweit möglich und sinnvoll, auf die vorliegende Evidenz der beschriebenen digitalen Gesundheitsanwendungen angewendet, um diese zu evaluieren. Abschließend werden zusammenfassend Besonderheiten in der Evaluation von digitalen Gesundheitsanwendungen beschrieben, sowie noch vorhandene Probleme und Limitationen der beschriebenen Frameworks genannt.

3.4.1 Anwendbarkeit der Frameworks

wenig Erfahrungswerte zur Anwendbarkeit

Zur generellen Anwendbarkeit von Assessment Frameworks für digitale Gesundheitsanwendung gibt es bislang wenig Erfahrungswerte [86]. Generell scheint es bei der Vielzahl an digitalen Gesundheitsanwendungen und der hohen Dynamik auf diesem Gebiet schier unmöglich, jede einzelne digitale Gesundheitsanwendung auf alle Aspekte zu evaluieren. Es empfiehlt sich also ein abgestuftes Vorgehen im Assessment von digitalen Gesundheitsanwendungen. Dieses abgestufte Vorgehen kann sich

abgestuftes Vorgehen in der Evaluation empfehlenswert

- zum ersten nach der Risikoklassifizierung oder dem Entwicklungsstadium der digitalen Gesundheitsanwendung richten. Eine Möglichkeit hierfür bietet das Evidence Standards Framework for Digital Health Technologies von NICE mit einer Einteilung nach Risikoklassen [28]. Eine weitere Möglichkeit für ein stufenhaftes Vorgehen bietet das mHealth Agile Development & Lifecycle Framework mit einer Einteilung nach Entwicklungsstadium [34].
- In einem zweiten Schritt können dann die relevanten HTA-Domänen und spezifische Aspekte für ein Assessment bestimmt und angewendet werden. Hier können etwa Frameworks wie das HTA Module for Medical Applications (MMAs) [27, 87] oder das DigiHTA [30] eingesetzt werden.

Für die in diesem Bericht ausgewählten digitalen Gesundheitsanwendungen wurde zur Überprüfung der Anwendbarkeit der Assessment Frameworks eben dieser abgestufte und kombinierte Ansatz gewählt: Dabei wurden die digitalen Gesundheitsanwendungen in einem ersten Schritt in Medizinproduktklassen nach 2017/745 MDR und nach den Evidenzstufen nach dem NICE Evidence Standards Framework eingeteilt und entsprechende geforderte Studiendesigns gegenübergestellt. In einem zweiten Schritt wurden mit Unterstützung der weiteren Frameworks (insb. die HTA MMA und die DigiHTA Checklisten) relevante Aspekte identifiziert, die für die jeweilige digitale Gesundheitsanwendung von besonders hoher Relevanz erscheinen.

definierte Studiendesigns für Entscheidungsgrundlage notwendig; Risikoeinteilung mit entsprechenden Studiendesigns; relevante Aspekte für jeweilige App

3.4.2 Zuteilung der ausgewählten digitalen Gesundheitsanwendung in Risikoklassen

Da für die nachfolgende Anwendung der Assessment Frameworks die Risikoklasse der digitalen Gesundheitsanwendung relevant ist, gibt die nachfolgende Tabelle eine Übersicht über die derzeitige Medizinproduktklasse (Angabe laut Hersteller), voraussichtliche Risikoklasse nach 2017/745 MDR sowie der funktionalen Klassifizierung nach Evidenzstufen nach dem NICE Evidence Standards Framework.

Einteilung zu Risikoklassen und funktionellen Evidenzstufen

An dieser Stelle soll erwähnt werden, dass die Risikoklassen nach 2017/745 und die Evidenzstufen nach NICE Ähnlichkeiten aufweisen, dennoch unterschiedlich definiert sind:

Tabelle 3.4-1: Risikoklassen nach 2017/745 MDR [18, 19] und Evidenzstufen nach NICE [28]

Risikoklassen nach EU MDR 2017/745 (Rule 11)	Evidenzstufen nach NICE
Risikoklasse 1 Jede Software, die nicht in Risikoklasse 2a oder höher eingeteilt wird	Stufe 1 Digitale Gesundheitsanwendungen ohne individuell messbare Effekte auf die Gesundheit von Anwender*innen, organisatorische Anwendungen auf Systemebene. <i>Beispiele:</i> elektronische Gesundheitsakte, Krankenhaus-Informationen-Systeme.
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	Stufe 2 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.

Risikoklassen nach EU MDR 2017/745 (Rule 11)	Evidenzstufen nach NICE
Risikoklasse 2b Software welche Information zur Entscheidung einer Diagnose oder therapeutischen Zwecken zur Verfügung stellt, bei der ernsthafte 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)	Stufe 3a 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.
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.	Stufe 3b 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.

Höherstufung der Risikoklasse für alle ausgewählten digitalen Gesundheitsanwendungen

Die meisten der hier beschriebenen digitalen Gesundheitsanwendungen befinden sich laut Hersteller in der Risikoklasse 1. Durch die neuen Risikoklassen nach 2017/745 MDR kommt es bei allen die für diesen Bericht ausgewählten digitalen Gesundheitsanwendungen zu einer Höherstufung der Risikoklasse. Nicht immer war eine eindeutige Zuteilung zu einer Risikoklasse möglich, jedoch befinden sich nach der neuen 2017/745 MDR alle für diesen Bericht ausgewählten digitalen Gesundheitsanwendung zumindest in der Risikoklasse 2a.

Tabelle 3.4-2: Einteilung der digitalen Gesundheitsanwendung nach derzeitiger Risikoklasse, Risikoklasse nach 2017/745 und Evidenzstufen nach NICE

	App	Derzeitige Medizinprodukt-Risikoklasse laut Hersteller	Voraussichtliche Risikoklasse nach 2017/745 MDR	Funktionelle Klassifizierung nach Evidenzstufen (NICE)
1.	SkinVision	k.A.	mind. 2a	3b
2.	Ada	1	mind. 2a	3b
3.	Mindable	1	mind. 2a	3b
4.	Natural Cycles	2b	2b	3b
5.	Preventicus Heartbeat	2a	2b?	3b
6.	DiabiLive	2b	2b	3b
7.	Kaia	1	2a	3b
8.	Moodpath	1	2a	3b
9.	Tinnitracks	1	2a	3b
10.	M-sense	1	2a	3b
11.	mySugr	1/2b	mind. 2b	3b

Risikoeinteilung einer DiGA nicht immer eindeutig

Dass die Risikoeinteilung einer digitalen Gesundheitsanwendung beziehungsweise eines Medizinprodukts – auch für Zulassungsbehörden – nicht immer eindeutig ist, zeigt sich auch am Beispiel der „iHealth Align Glucose Meter and Reveal LINQ Insertable Cardiac Monitoring System“. Dieses Medizinprodukt wird von der australischen Therapeutic Goods Administration (TGA) und der US-amerikanischen Food and Drug Administration (FDA) unterschiedlich klassifiziert [26].

Da die in diesem Bericht ausgewählten digitalen Gesundheitsanwendungen nach der funktionellen Klassifizierung nach Evidenzstufen (NICE) als Evidenzstufe 3b einzuteilen sind, ergeben sich

- Mindestanforderung für hochwertige vergleichende Interventionsstudien mit relevanten Endpunkten (z. B. Diagnostische Genauigkeit, Patient-relevante Endpunkte wie Lebensqualität und Nutzerzufriedenheit, Surrogatendpunkte wie physiologische Messwerte) bis
- Maximalanforderung („Best practice“) hochwertige randomisierte kontrollierte Studien, welche die digitale Gesundheitsanwendung im jeweiligen Setting mit dem derzeitigen Versorgungsstand vergleicht. Dabei sollte ein Nutznachweis für die Anwender*innen in den für das Anwendungsgebiet relevanten Endpunkten gezeigt werden.

**ausgewählte DiGA
entsprechen NICE
Evidenzstufe 3b**

3.4.3 Anwendung der HTA-Assessment Frameworks

Die Evaluierung von digitalen Gesundheitsanwendungen unterscheidet sich von herkömmlichen HTA-Methoden, da zusätzliche Domänen wie Datenschutz oder Konnektivität berücksichtigt werden müssen. Auch bei der Bewertung der Effektivität und Sicherheit gibt es Unterschiede im Vergleich zu konventionellen Therapien.

zusätzliche Domänen

HTA Module for Mobile Medical Applications (MMA) [12, 26, 87]

Dieses Framework soll komplementär zum Standard HTA Framework angewendet werden. In der praktischen Anwendung zeigt das Framework Limitationen: für jede HTA Domäne werden für die digitale Gesundheitsanwendung zusätzliche Überlegungen genannt, ohne jedoch Bewertungskriterien für Antworten oder Risikoklassifizierungen zu definieren.

**keine
Risikoklassifizierung**

Evidence Standards Framework for Digital Health Technologies [28]

Dieses Framework sollte laut Autor*innen nur für Technologien welche fixierte Algorithmen nutzen verwendet werden. Trotz dieser Limitation zeigt sich das Framework in der praktischen Anwendung als sinnvoll: Mit Hilfe dieses Frameworks wird eine digitale Gesundheitsanwendung zunächst einer funktionellen Klassifizierung nach Evidenzstufen unterzogen. Digitale Gesundheitsanwendung, die trotz hoher Klassifizierung der Evidenzstufe 3b keinerlei Studien mit Komparator (bestenfalls randomisiert kontrollierte Studien) vorweisen können, können mit Hilfe dieses Frameworks eindeutig von einer Refundierung ausgeschlossen werden. Für digitale Gesundheitsanwendungen, deren Studiendesigns den Anforderungen der jeweiligen Evidenzstufen entsprechen, empfiehlt sich eine weiterführende Evaluation.

**Risikoklassifizierung
über Evidenzstufen,
keine Anwendung bei
adaptiven Algorithmen**

Tabelle 3.4-3: Potentielle Empfehlung ausgewählter digitaler Gesundheitsanwendungen unter Anwendung von [28] basierend auf vorliegender Evidenz und relevante Bewertungsaspekte

	App	Potentielle Empfehlung für weitere Evaluierung unter Anwendung von NICE Framework [28]	Begründung	Besonders relevante Aspekte nach Assessment Frameworks [27, 30, 87]
1.	SkinVision	Nein	Beobachtungsstudien: Fall-Kontrollstudien vorhanden Interventionsstudien: keine vorhanden	Wirksamkeit: diagnostische Genauigkeit Sicherheit: Fehlinformation Kosteneffektivität: vermiedene/Mehrkosten Artificial Intelligence: Nachvollziehbarkeit des Algorithmus Organisatorische Auswirkungen: Überdiagnostik Datenschutz und Haftung: gesundheitsbezogene Daten, Haftung bei Fehldiagnose Ethik: Ver(un)sicherung
2.	Ada	Nein	Beobachtungsstudie: Fall-Kontrollstudie vorhanden Interventionsstudien: keine vorhanden	Wirksamkeit: diagnostische Genauigkeit Sicherheit: Fehlinformation Kosteneffektivität: vermiedene/Mehrkosten Artificial Intelligence: Nachvollziehbarkeit des Algorithmus Organisatorische Auswirkungen: Überdiagnostik Datenschutz und Haftung: gesundheitsbezogene Daten, Haftung bei Fehldiagnose Ethik: Ver(un)sicherung
3.	Mindable	Nein (Re-Evaluierung bei Vorliegen der Studienergebnisse)	Laufende Studien ohne genanntes Studiendesign	Wirksamkeit: Reduktion von Panikstörung und Agoraphobie; Sicherheit: unerwünschte Nebenwirkungen Kosteneffektivität: vermiedene/Mehrkosten Artificial Intelligence: Nachvollziehbarkeit des Algorithmus Organisatorische Auswirkungen: Interoperabilität Datenschutz und Haftung: gesundheitsbezogene Daten, Haftung bei unerwünschten Nebenwirkungen Ethik: Ver(un)sicherung
4.	Natural Cycles	Nein	Beobachtungsstudien: vorhanden Interventionsstudien: keine vorhanden	Wirksamkeit: Bestimmung des Eisprungs bei Kinderwunsch, oder Empfängnisverhütung; Sicherheit: ungewollte Schwangerschaft, Aborte Kosteneffektivität: reduzierte oder erhöhte Folgekosten Artificial Intelligence: Nachvollziehbarkeit des Algorithmus Organisatorische Auswirkungen: Interoperabilität Datenschutz und Haftung: gesundheitsbezogene Daten, Haftung bei ungewollter Schwangerschaft Ethik: Ver(un)sicherung
5.	Preventicus Heartbeat	Nein	Beobachtungsstudien: Fall-Kontrollstudien vorhanden Interventionsstudien: keine vorhanden	Wirksamkeit: diagnostische Genauigkeit Sicherheit: Fehlinformation Kosteneffektivität: vermiedene/Mehrkosten Artificial Intelligence: Nachvollziehbarkeit des Algorithmus Organisatorische Auswirkungen: Überdiagnostik, Interoperabilität Datenschutz und Haftung: gesundheitsbezogene Daten, Haftung bei Fehldiagnose Ethik: Ver(un)sicherung
6.	DiabiLive	Nein	Keine Studien vorhanden	Wirksamkeit: diagnostische Genauigkeit, Management von Blutzuckerwerten Sicherheit: Fehlinformation, unerwünschte Nebenwirkungen Kosteneffektivität: vermiedene/Mehrkosten Artificial Intelligence: Nachvollziehbarkeit des Algorithmus Organisatorische Auswirkungen: Überdiagnostik, Interoperabilität Datenschutz und Haftung: gesundheitsbezogene Daten, Haftung bei unerwünschten Nebenwirkungen Ethik: Ver(un)sicherung

	App	Potentielle Empfehlung für weitere Evaluierung unter Anwendung von NICE Framework [28]	Begründung	Besonders relevante Aspekte nach Assessment Frameworks [27, 30, 87]
7.	Kaia	Eventuell (nach kritischer Evaluierung der RCTs)	Beobachtungsstudien: Kohortenstudien vorhanden Interventionsstudien: RCTs vorhanden	Wirksamkeit: diagnostische Genauigkeit, Reduktion von Schmerzsymptomatik Sicherheit: Fehlinformation, unerwünschte Nebenwirkungen Kosteneffektivität: vermiedene/Mehrkosten Artificial Intelligence: Nachvollziehbarkeit des Algorithmus Organisatorische Auswirkungen: Überdiagnostik, Interoperabilität Datenschutz und Haftung: gesundheitsbezogene Daten (Video-Input), Haftung bei unerwünschten Nebenwirkungen Ethik: Ver(un)sicherung
8.	Moodpath	Nein	Beobachtungsstudien: Studie zur Nutzung der Anwendung vorhanden Interventionsstudien: keine vorhanden	Wirksamkeit: diagnostische Genauigkeit, Reduktion depressiver Symptomatik Sicherheit: Fehlinformation, unerwünschte Nebenwirkungen Kosteneffektivität: vermiedene/Mehrkosten Artificial Intelligence: Nachvollziehbarkeit des Algorithmus Organisatorische Auswirkungen: Überdiagnostik Datenschutz und Haftung: gesundheitsbezogene Daten, Haftung bei Fehldiagnostik oder unerwünschten Nebenwirkungen Ethik: Ver(un)sicherung
9.	Tinnitracks	Eventuell (nach kritischer Evaluierung der RCTs)	Beobachtungsstudien: keine vorhanden Interventionsstudien: RCTs zu TMNMT vorhanden, RCT zu Tinnitracks nicht vorhanden	Wirksamkeit: Reduktion von Tinnitus-Symptomatik Sicherheit: unerwünschte Nebenwirkungen Kosteneffektivität: vermiedene/Mehrkosten Artificial Intelligence: Nachvollziehbarkeit des Algorithmus Organisatorische Auswirkungen: Interoperabilität Datenschutz und Haftung: gesundheitsbezogene Daten, Haftung bei unerwünschten Nebenwirkungen Ethik: Ver(un)sicherung
10.	M-sense	Nein (Evaluierung bei Vorliegen der Studienergebnisse)	Laufende Studien (RCT)	Wirksamkeit: diagnostische Genauigkeit, Reduktion von Kopfschmerzsymptomatik Sicherheit: Fehlinformation, unerwünschte Nebenwirkungen Kosteneffektivität: vermiedene/Mehrkosten Artificial Intelligence: Nachvollziehbarkeit des Algorithmus Organisatorische Auswirkungen: Überdiagnostik Datenschutz und Haftung: gesundheitsbezogene Daten, Haftung bei unerwünschten Nebenwirkungen Ethik: Ver(un)sicherung
11.	mySugr	Eventuell (nach kritischer Evaluierung der RCTs) (Re-Evaluierung bei Vorliegen der Studienergebnisse)	Beobachtungsstudien: vorhanden Interventionsstudien: 1 RCT Cross-Over vorhanden Aktuell laufende RCTs	Wirksamkeit: diagnostische Genauigkeit, Management von Blutzuckerwerten Sicherheit: Fehlinformation, unerwünschte Nebenwirkungen Kosteneffektivität: vermiedene/Mehrkosten Artificial Intelligence: Nachvollziehbarkeit des Algorithmus Organisatorische Auswirkungen: Überdiagnostik, Interoperabilität Datenschutz und Haftung: gesundheitsbezogene Daten, Haftung bei unerwünschten Nebenwirkungen Ethik: Ver(un)sicherung

	Digital Healthcare Services/Digi-HTA [30]
keine Bewertungskriterien oder Risikoklassen	Dieses Framework soll komplementär zu bisher verfügbaren HTA-Methoden angewendet werden. Darunter finden sich Domänen, die für digitale Gesundheitsanwendungen spezifisch sind. In der praktischen Anwendung zeigt das Framework Limitationen: Für jede Domäne werden Fragen gestellt, z. B. „Wie holt sich der Hersteller Feedback der Anwender*innen und passt die digitale Gesundheitsanwendung entsprechend diesem Feedback an?“, ohne jedoch Bewertungskriterien für Antworten oder Risikoklassifizierungen zu definieren.
	Design & Evaluation Framework for Digital Health Interventions/DEDHI [32]
keine Risikoklassen; Phasen der Produktentwicklung	Dieses Framework fordert Evidenz je nach Produktentwicklungsstadium der digitalen Gesundheitsanwendung. Während in der frühen Entwicklungsphase Studien zu Durchführbarkeit gefordert werden, werden in der späteren Entwicklungsphase randomisiert kontrollierte Studien gefordert sowie kontinuierliche Erhebungen zu Reichweite, Impact und Nebenwirkungen.
	Digital Health Scorecard [33]
keine Scores definiert	Dieses Framework fordert Evidenz für zuvor definierte klinische Endpunkte. Dabei soll ein Vergleich zum bisher existierenden Versorgungsstandard gezogen werden. Dieses Framework macht keine Angaben zu vorgeschlagenen Studiendesigns. In der derzeitigen Form zeigt dieses Framework keine ausreichende Anwendbarkeit, bei zukünftiger Definition von Scores zu den einzelnen Kriterien des Frameworks kann sich die Anwendbarkeit verbessern.
	mHealth Agile Development & Lifecycle [34]
keine Risikoklassen; Phasen der Produktentwicklung	Dieses Framework besteht aus kontextuellen Fragen zu den digitalen Gesundheitsanwendungen. Dabei wird das Entwicklungsstadium in den Evidenzerfordernissen berücksichtigt. In der frühen Phase fordert das Framework Kohortenstudien und Fallserien, in den späteren Phasen randomisiert kontrollierte Studien und Post-Market-Surveillance. Dieses Framework unterstützt bei der Zuordnung von Studiendesigns zu Produktentwicklungsphasen, macht dabei jedoch keine präzise Einteilung nach Risikoklasse der digitalen Gesundheitsanwendung.

3.4.4 Zusammenfassung und Diskussion der Anwendung der Assessment Frameworks

Mindestvoraussetzung für Studiendesigns notwendig	Eine gute Anwendbarkeit eines Assessment Frameworks setzt eine bestmögliche Grundlage zur Entscheidungsfindung bei Refundierungsentscheidungen voraus, ist aber auch die Voraussetzung, dass mHealth Hersteller eine Orientierung haben, welche Evidenz sie vorlegen müssen. Dazu bedarf es klarer Mindestvoraussetzungen für gewünschte Studiendesigns. Frameworks, welche keine Angaben zum gewünschten Studiendesign angeben und ausschließlich deskriptive Fragen zur digitalen Gesundheitsanwendung auflisten, können zwar Unterstützung bei der Entscheidungsfindung bieten, jedoch keine zufriedenstellende Entscheidungsgrundlage für evidenzbasierte Entscheidungen in Bezug auf Refundierungen. Nur wenige der identifizierten Frameworks können diese Anforderung erfüllen und sind ausreichend konkret in der Anwendung, um eine Anleitung zu geben.
auch zur Orientierung für Hersteller	

Neben der Evidenzklassifizierung gibt es weitere Aspekte, die ein Assessment Framework hilfreich machen könnten, wie etwa

- die Verwendung des PICO-Schema zur besseren Festlegung der Intentionen (Endpunkte) der untersuchten Intervention bei definierten Anwender*innen-Gruppen (Patientenpopulation), oder
- eine Taxonomie der Interventionen zur Klassifizierung der mHealth Anwendungen insb. zur Vergleichbarkeit bei gleichen Indikationen/Patientenpopulationen, oder
- die Verwendung von (gleichen) Fragenkatalogen bei mHealth Anwendungen mit gleichen Funktionen im Management von (auch unterschiedlichen) Erkrankungen.

Ein Problem in der Anwendung ist der unterschiedliche Anwendungsbereich der hier beschriebenen Frameworks. Beispielsweise können das HTA Module for Mobile Medical Applications (MMA) [27] oder das Digi-HTA Framework für mHealth [30] uneingeschränkt für alle in diesem Bericht ausgewählten digitalen Gesundheitsanwendungen eingesetzt werden. Das NICE Evidence Standards Framework für Digital Health Technologies stößt hier mitunter schon an Grenzen [28]: so empfehlen die Autor*innen des Frameworks keine Anwendung auf digitale Gesundheitsanwendungen, die künstliche Intelligenz mit adaptiven Algorithmen verwenden. Für die in diesem Bericht beschriebenen digitalen Gesundheitsanwendungen ist die genaue Funktionsweise des Algorithmus nicht bekannt, adaptive Algorithmen können nicht ausgeschlossen werden. Trotz dieser Limitationen wird das NICE Evidence Standards Framework von vielen Institutionen als das derzeit sinnvollste Assessment Framework erachtet (siehe INAHTA ListServ Antworten im Appendix 7.2.1).

Eine kontinuierliche Bewertung im Sinne eines Monitoring der digitalen Gesundheitsanwendung ist anzustreben. So können Aktualisierungen, insbesondere grundsätzliche Änderungen des verwendeten Algorithmus oder das Integrieren neuerer Funktionsweisen der digitalen Gesundheitsanwendung bewertet werden. Es ist in den hier beschriebenen Assessment Frameworks jedoch nicht definiert, ab wann ein Re-Assessments notwendig ist. Digitale Gesundheitsanwendungen unterliegen häufigen, mitunter wöchentlichen Updates. Nicht jedes Update bedarf eines umfassenden erneuten Assessments. Zur Veranschaulichung: ein Update des User Interface zur Verbesserung der Benutzerfreundlichkeit bedarf wahrscheinlich keiner weiteren Evaluierung, ein Update zur Anpassung eines Algorithmus, um höhere Sensitivität und Spezifität in der Diagnostik zu erreichen bedarf möglicherweise einer weiteren Teil-Evaluierung. Das Hinzufügen von diagnostischen Funktionen, wenn die digitale Gesundheitsanwendung zuvor rein informative und therapeutische Funktionen beinhaltete, bedarf einer weiteren Evaluierung. Wenn die digitale Gesundheitsanwendung durch ein Update in eine höhere Risikoklasse aufsteigt bedarf es jedenfalls einer gründlichen Re-Evaluierung.

Bei der Anwendung der Assessment Frameworks empfiehlt sich zunächst eine Einteilung nach Risikoklasse und infolge eine Auswahl an relevanten Bewertungsaspekten, die sich an den (geweckten) Erwartungen zum Potential und Befürchtungen zu Folgewirkungen orientieren, zu treffen, statt allumfassende Assessments durchzuführen.

weitere Aspekte, die Assessment Framework nützlich machen

unterschiedliche Anwendungsbereiche, Limitationen bei adaptiven Algorithmen

Re-Evaluierung je nach Umfang des Software-Updates

Einteilung in Risikoklassen, relevante Bewertungsaspekte

3.5 Übersicht von Regelungen in ausgewählten Ländern

USA: Digital Health Innovation Action Plan	In den USA wurde der „Digital Health Innovation Action Plan“ ins Leben gerufen. In diesem ist ein „Precertification“-Programm vorgesehen, durch welches in der Zertifizierung der Fokus nicht auf einzelnen digitalen Gesundheitsanwendungen liegen soll, sondern die Zertifizierung auf Ebene von Entwicklern und Prozessen stattfindet: digitale Gesundheitsanwendungen, die von zertifizierten Entwicklern hergestellt wurden, wären nach dem „Digital Health Innovation Action Plan“ zertifiziert, ohne dass es einer Bewertung der einzelnen digitalen Gesundheitsanwendung bedarf [88].
Zertifizierung von Entwickler*innen	
Europa: Refundierung einzelner DiGA	In Europa hat man in einzelnen Ländern begonnen, einen Prozess der Registrierung aufzusetzen, der infolge für einzelne digitale Gesundheitsanwendungen auch zu einer Refundierung durch Sozial- und Krankenversicherungen führen kann. Im Folgenden werden einzelne Länder und ihr Vorgehen kurz beschrieben.

3.5.1 Belgien

Validierungspyramide für DiGA	Der administrative Rahmen für die Nutzung von Gesundheitsapplikationen wird auf der Plattform mHealthBelgium (https://mhealthbelgium.be/) vorgestellt. Zentraler Bestandteil des Rahmenwerks ist die so genannte „Validierungspyramide“ für Health Apps [24].
zentrale Plattform für Informationen zu DiGA	Diese Plattform zentralisiert alle relevanten und erforderlichen Informationen über mobile Apps für Patient*innen, Angehörige der Gesundheitsberufe und Gesundheitseinrichtungen in drei Sprachen. Die Informationen beziehen sich auf die CE-Kennzeichnung, den Datenschutz, die Kommunikationssicherheit, die Interoperabilität mit anderen IT-Systemen und die Art und Weise, wie die App finanziert wird.
3 Ebenen	Die Validierungspyramide besteht aus drei Ebenen; eine App tritt immer auf der unteren Ebene ein und kann in der Hierarchie aufsteigen [89]: ■ <i>Stufe 1 (M1)</i> – die Anwendung ist ein CE-zertifiziertes Medizinprodukt; freiwillige Meldung bei der belgischen Federal Agency for Medicines and Health Products (FAMHP), wo die CE-Zertifizierung sowie die Einhaltung der EU-Datenschutzverordnung (GDPR) überprüft und bestätigt wird. ■ <i>Stufe 2 (M2)</i> – die Anwendung ist ein CE-zertifiziertes Medizinprodukt und verfügt über Interoperabilität und Konnektivität mit den nationalen eHealth Anforderungen. Zusätzlich muss die digitale Anwendung eine Sicherheits-Risikobewertung (von einer unabhängigen Organisation entwickelt) vorlegen, um in mHealthBelgium aufgenommen zu werden. ■ <i>Stufe 3 (M3)</i> – die Anwendung ist ein CE-zertifiziertes Medizinprodukt, ist nachweislich sicher und ein sozio-ökonomischer Nutzen ist nachgewiesen. Sie verfügt über einen Erstattungscode von INAMI/RIZIV.

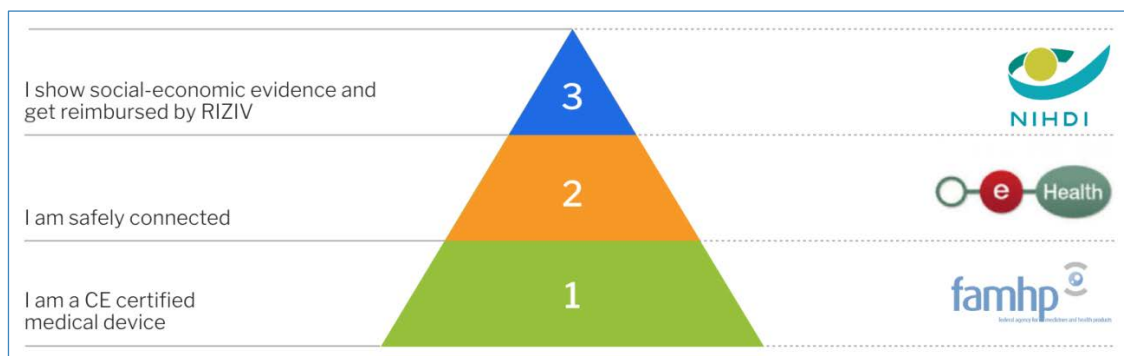


Abbildung 3.5-1: Belgische Validierungspyramide (<https://mhealthbelgium.be/validation-pyramid>)

Die Plattform ist eine Initiative von drei Institutionen und ging im Jänner 2019 online [89]:

1. FAMHP (Bundesamt für Arzneimittel und Gesundheitsprodukte): Die Zulassungsbehörde ist innerhalb von mHealthBelgium für die Ebene M1 zuständig.
2. eHealth-Plattform (eine Einrichtung der Bundesregierung) hat den Auftrag, elektronische Dienste und Datenaustausch zwischen allen Beteiligten im Gesundheitswesen zu fördern und zu unterstützen. eHealth ist für die Ebene M2 verantwortlich.
3. NIHDI (Nationales Institut für Gesundheits- und Behindertenversicherung) ist für die Rückerstattung von Medikamenten, medizinischen Geräten und medizinischen Vorräten zuständig. NIHDI ist für die Ebene M3 zuständig.

**Kooperation von
3 Institutionen**

Im September 2020 waren 22 Gesundheitsapplikationen auf der Plattform mHealth Belgium vertreten (<https://mhealthbelgium.be/apps>): 17 auf Stufe 1, 5 auf Stufe 2 und keine auf Stufe 3 der Validierungspyramide.

**Plattform mit 22 DiGA,
noch keine auf Stufe 3**

Derzeit gibt es in Belgien keinen nationalen Erstattungsrahmen für Gesundheitsapplikationen. NIHDI und INAMI/RIZIV arbeiten daran: Der Zeitpunkt für die Veröffentlichung des neuen Rahmens ist nicht klar [89]. In Ermangelung einer nationalen Rückerstattung können einzelne Versicherungsgesellschaften ihren Versicherten die Abgeltung für Gesundheitsapplikationen anbieten.

**nationale Refundierung
in Arbeit**

Seit Januar 2020 gibt es mindestens zwei unabhängige Krankenkassen (Partenamut & Partena Ziekenfondsen) [24]:

- *Partenamut* hat die Rückerstattung in groben Zügen definiert; die Versicherung bietet eine feste Rückerstattung von bis zu 20 € pro Begünstigten pro Jahr. Das Erstattungsmodell deckt also nicht alle Apps der Plattform mHealthBelgium gleichzeitig ab. Nur eine begrenzte Anzahl von Gesundheitsapplikationen kann innerhalb des angegebenen Erstattungsbetrags abgedeckt werden. Voraussetzung ist, dass die App vom Arzt verschrieben oder von Partenamut anerkannt ist.
- *Partena Ziekenfondsen* (PZ) verfolgt dagegen einen anderen Ansatz. Drei Mal pro Jahr leistet PZ einen Beitrag von 10 € pro Konsultation (insgesamt 30 € pro Jahr) für Präventionsdienste (Krebs, Herz-Kreislauf-Erkrankungen, Allergien, Osteoporose, Atemwegserkrankungen und Diabetes). Dazu gehören insbesondere Gesundheitsapplikationen: Fibrichk-App kostet 25 € pro Monat. PZ sieht daher einen Zuschuss

**Selektivverträge mit
2 unabhängigen
Krankenkassen**

von 30 € pro Jahr vor; PZ bietet einen Zuschuss von 30 € pro Jahr für die moveUp Coach App (Reha bei Hüft-/Knieoperationen); PZ bietet 50 % Rabatt-Codes für SkinVision an (die Patient*innen zahlen 15 € statt 30 € für das Jahresabonnement);

aktuell laufende klinische Studie

Für die digitale Anwendung moveUp Coach (Stufe 2) ist derzeit eine klinische Studie und Kontrahierung zwischen INAMI/RIZIV und 15 Spitälern in Planung. Diese Studie wird von der Universität Gent in Kooperation mit dem belgischen HTA-Institut KCE durchgeführt und vergleicht ein klassisches Rehabilitationsprogramm mit einem, das durch die App moveUp Coach unterstützt wird. Die Studie soll im Herbst 2021 eine Entscheidung bezüglich einer Refundierung durch INAMI/RIZIV unterstützen [90].

3.5.2 Frankreich

zentrale LPPR-Liste für Erstattungen

August 2020: erste DiGA in Erstattungsliste aufgenommen

Entsprechend dem vom nationalen HTA-Institut HAS veröffentlichten „Leitfaden zu den Besonderheiten der klinischen Bewertung eines verbundenen Medizinproduktes (CMD) im Hinblick auf seinen Erstattungsantrag“ [91] sollte das wichtigste Erstattungsinstrument für Gesundheitsapplikationen die „Liste des Produits et Prestations Remboursables“ (LPPR-Liste) sein. Der Prozess der Aufnahme einer Gesundheitsanwendung in die LPPR-Liste ähnelt dem bekannten Weg für implantierbare Geräte, invasive nicht implantierbare Geräte und medizinische Hilfsmittel. Im August 2020 wurde die erste digitale Gesundheitsanwendung (Moovecare poumon) in die LPPR-Liste zur Erstattung aufgenommen [92]. „Moovecare poumon“ wird zur telemedizinischen Nachsorge von Lungenkrebs-Rezidiven und Komplikationen bei Patient*innen mit Lungenkrebs eingesetzt. Die App wurde mit einem Zusatznutzen Level 3 (moderate Verbesserung im Vergleich zu herkömmlichem Monitoring, Risikoklasse 1) eingestuft. Bislang wurden zwei Anwendungen von CNEDiMTS, dem Nationalen Komitee zur Evaluierung von Medizinprodukten und Gesundheitstechnologien evaluiert: Diabeo (App, zur Unterstützung von Typ-1-Diabetes Patient*innen bei der Dosierung von Insulin, Risikoklasse 2b) und Moovecare poumon [92].

seltener: Erstattung der DiGA über die Kosten des Implantats

Ein weiterer (aber bislang sehr seltener) Ansatz ist – etwa bei Telemonitoring von Herzschrittmacher-Implantierten oder anderen teuren Implantaten –, die erstattungsfähigen Kosten für die Fernüberwachung bereits in den Kosten des Implantats zu inkludieren: als Beispiel gilt der LPPR-gelistete Einkammer-Herzschrittmacher Azure XT (Medtronic), der zwei Erstattungsoptionen hat: Ein (Hardware-)Sender – MYCARELINK – und eine „Gesundheits-App-Sender“ MYCARELINK HEART [24].

ETAPES-Programm: Förderung von DiGA, Evidenzgenerierung

Zusätzlich zu dem zentralisierten Weg der LPPR-Eintragung (der noch in Umsetzung ist) werden digitale Gesundheitsanwendungen derzeit auch über das experimentelle, nur temporäre und öffentlich geförderte Programm für Telemonitoring (ETAPES, *Expérimentation de Télémédecine pour l'Amélioration des Parcours en Santé*) erstattet und können Gegenstand von Programmen zur Generierung von Evidenz sein. ETAPES umfasst fünf klinische Bereiche für chronische Patient*innen [24]:

- Herzinsuffizienz,
- Niereninsuffizienz,
- Atemwegsinsuffizienz,
- Diabetes und
- implantierbare Medizinprodukte bei chronischen Herzerkrankungen

Die Finanzierung des ETAPES-Programms umfasst drei Komponenten [24]:

- die Bezahlung des/der Arztes/Ärztin, der/die das Telemonitoring durchführt,
- die Bezahlung des medizinischen Fachpersonals, das die therapeutische Unterstützung der Patient*innen leistet, und
- die Bezahlung des Anbieters einer technischen Lösung für das Telemonitoring. Die technische Lösung für das Telemonitoring kann ein angeschlossenes medizinisches Gerät, eine digitale Gesundheitsanwendung, eine digitale Plattform oder eine Kombination aller oben genannten Elemente umfassen.

**3 Aspekte
der Finanzierung**

Darüber hinaus gibt es an zwei an Spitälern angesiedelte Förderungsprogramme für „coverage with evidence development (CED)“: Program for Hospital Clinical Research (PHRC) and the Program for Medical Economic Research (PRME), die nicht nur, aber auch für digitale Gesundheitsanwendungen eingesetzt werden [93].

**weitere Förderungen
für DiGA**

Es ist geplant, dass die Bewertung von digitalen Gesundheitstechnologien vor einer Kostenerstattung einem HTA durch HAS, der nationalen HTA-Institution, unterzogen wird. Die Bewertungen basieren auf zwei Methodenpapieren von 2016 und 2018. Gegenwärtig sind die Möglichkeiten für die Erstattung von digitalen Gesundheitsapplikationen in Frankreich begrenzt; die Einforderung von Nachweisen auf hohem Evidenzniveau ist geplant [91, 94].

**Bewertung durch
nationales HTA-Institut,
hohe Evidenzerfordernisse**

3.5.3 Niederlande

Der *Raad voor Volksgezondheid en Samenleving* (RVS, Council of Public Health & Society: (<https://www.raadvr.nl/about-the-rvs>)), ein unabhängiges, die Regierung beratendes Gremium, legte die definitorische Grundlage – Klassifikationen – für eine mögliche Bewertung und Rückerstattung von digitalen Anwendungen fest [95]:

**Klassifikation
als Grundlage für
Rückerstattung**

- *e-Care*: e-Diagnose, e-Konsultationen, e-Care wie z. B. Monitoring, e-Präventionsintervention mit hohem Risiko bei einer Person;
- *E-Support*: E-Zugang zu Patientenakten, E-Management wie z. B. Online-Terminvereinbarung;
- *e-Gesundheit*: e-Gesundheitserziehung, e-Prävention wie z. B. Erkennung bestimmter Risikogruppen.

Diese Klassifikation des Begriff e-Health wurde von den zwei zentralen Institutionen bei der Einführung von digitalen Gesundheitsanwendungen – vom *Centre of Expertise for Standardization and eHealth* (Nictiz: <https://www.nictiz.nl/>), sowie vom *National eHealth Living Lab* (NeLL: <https://nell.eu>) – übernommen. Mit der Einrichtung des *National eHealth Living Lab* (NeLL) im Jahr 2018 hat das niederländische Gesundheitsministerium die ersten Schritte zu einem einheitlichen nationalen Apps-Archiv unternommen.

**2 zentrale Institutionen:
Nictiz und NeLL**

Das NeLL ist ein unabhängiges wissenschaftliches Netzwerk, das mit verschiedenen Partnern (dem niederländischen Gesundheitsministerium, verschiedenen Medizintechnik-Herstellern und akademischen Institutionen) zusammenarbeitet. Ziel ist es, effektive wissens-basierte Werkzeuge für den Einsatz von eHealth innerhalb bestehender Versorgungsprozesse zur Verfügung zu stellen. NeLL wurde eingerichtet, um die Apps zu bewerten und ein Repository zu entwickeln, welche Apps effizient, effektiv und einfach zu benutzen sind. Kriterien zur Bewertung sind [96]:

**NeLL: Verzeichnis für
effiziente, effektive und
nutzerfreundliche DiGA**

	■ <i>Anwendernutzen:</i> Die Endbenutzer werden zu jeder Zeit in die Entwicklung von Ideen, Konzepten, Produkten und Dienstleistungen einbezogen, um sicherzustellen, dass die entwickelte Anwendung dem Bedarf entspricht und eine positive Wirkung hat. ■ <i>Nachgewiesene Wirkung:</i> Das eHealth-Tool wird durch wissenschaftliche Forschung unterstützt und hat nachweislich Wirkung gezeigt. ■ <i>Transparenz:</i> es erfolgt eine transparente Kommunikation über den Fortschritt aller NeLL-Projekte und die wissenschaftliche Forschung auf der NeLL-Website, in sozialen Medien und in Publikationen.
14 DiGA, noch keine NeLL-kompatibel	Seit Februar 2020 gibt es 14 Anwendungen auf der NeLL-Plattform, von denen keine die Kennzeichnung „NeLL-kompatibel“ trägt [24].
bisher keine nationale Refundierung	Derzeit gibt es keine offiziellen national-gültigen administrativen Rahmenbedingungen für den Einsatz und die Refundierung von digitalen Gesundheitsapplikationen. Es ist im Verantwortungsbereich der einzelnen Krankenkassen zu entscheiden, ob sie ihren Versicherten Gesundheitsapplikationen zur Verfügung stellen. Das Kostenerstattungsmodell für Gesundheitsapplikationen unterscheidet sich je nach dem Setting, in dem sie eingesetzt wird: Nach Vorgaben der niederländischen Gesundheitsbehörde (NZa) fallen Gesundheitstechnologien, die mit der fachärztlichen Versorgung im Krankenhaus verbunden sind, in der Regel unter die Erstattung über DRG. Im Gegensatz dazu werden digitale Anwendungen, die für den Einsatz im niedergelassenen Bereich/im häuslichen Umfeld bestimmt sind, durch die individuelle Entscheidung der Versicherungsgesellschaften erstattet [24].
Setting: Versorgung mit DiGA im Krankenhaus über DRG, im ambulanten Bereich Selektivverträge	
Versorgungsprozesse und -kosten, Patientenergebnisse	Hauptkriterien für Krankenkassen sind die Verbesserung der Versorgungsprozesse, der Patientenergebnisse und die Senkung der Versorgungskosten. Sobald Daten und Informationen zur Wirksamkeit des Einsatzes der Technologie vorliegen, kann eine Krankenversicherung einen Vertrag schließen. Theoretisch ist es auch möglich, das Versicherungsbudget für Anwendungsstudien (coverage with evidence development) zu nutzen, um Nachweise zu generieren, aber die Anzahl der im Rahmen des Programms vergebenen Projekte ist sehr gering.
Selektivverträge mit SkinVision und Luscii	Beispiele für erstattete Gesundheitsapplikationen sind SkinVision und Luscii, die aber nicht die einzigen beiden Gesundheitsapplikationen, die in den Niederlanden von den Krankenkassen übernommen werden, sind [24]: ■ SkinVision und CZ Healthcare Insurance haben 2018 mit einem Pilotprogramm in Kooperation mit niederländischen Ärzt*innen zur Datengenerierung begonnen und die Applikation wurde inzwischen in den Leistungskatalog von CZ aufgenommen. ■ Über die Luscii-App bleiben der/die Arzt/Ärztin und die Patient*innen (mit COPD, Herzinsuffizienz) in Kontakt, tauschen Informationen aus, und der Arzt kann den Patient*innen Leitlinien zur Verfügung stellen. Luscii wird von 94 % der niederländischen Versicherungen erstattet.
eigene Abrechnung für telemedizinische Leistungen	Für telemedizinische Dienste gelten eigene Abrechnungspositionen: Im Jahr 2021 wird die Leistung „Telefonische Beratung“ durch die neue Leistung „Triage-Beratung“ ersetzt: Es wird zwischen dem Erstkontakt (Triage) und der möglichen Nachbetreuung (Beratung oder Besuch) unterschieden. Auch die Kostenerstattung für Technologien zur häuslichen Pflege ist seit 2020 erweitert: der Leistungserbringer kann maximal 4 Stunden Bildschirm-Kommunikation und maximal 2,5 Stunden pharmazeutische Fernbetreuung pro Patient*in und Monat zum vereinbarten Stundensatz für Pflege- und persönliche Betreuungsleistungen in Anspruch nehmen.

3.5.4 England

Die wichtigste administrative Voraussetzung für eine öffentliche Unterstützung ist die Registrierung einer digitalen Gesundheitsanwendung in der NHS Apps Library (<https://www.nhs.uk/apps-library/>), die vom NHS Digital verwaltet wird. Die NHS Apps Library sammelt alle Gesundheits-Apps, die anhand nationaler Standards bewertet wurden und die nachweislich sicher sind. Die Aufnahme der Apps in der NHS Apps Library bedeutet jedoch keine Rückerstattung, da dies individuell von den Clinical Commissioning Groups (CCGs) in den National Health Service (NHS) Trusts entschieden wird [24]. Die NHS Apps Library erleichtert den CCGs eine Entscheidung über den Einsatz der App. Im Februar 2020 gab es 79 Apps, im September 2020 bereits 97 Apps in der NHS Apps Library.

Die Apps in der NHS Apps Library werden anhand von Standards, die von *NHSX Digital Health Technology* (<https://www.nhsx.nhs.uk/>) entwickelt wurden, bewertet. NHSX ist eine organisationsübergreifende Einheit, die Teams des Gesundheits- und Sozialministeriums, des NHS England und des NHS Improvement zusammenbringt, um an der digitalen Transformation des NHS zu arbeiten. Die NHSX Standards für digitale Gesundheitstechnologien basieren auf bestehenden Beurteilungskriterien wie Wirksamkeit, Sicherheit, Datenschutz, Robustheit, Stabilität, Interoperabilität, Benutzerfreundlichkeit, Zugänglichkeit und Verantwortlichkeit.

Das im März 2019 vom NICE veröffentlichte „Evidence standards framework for digital health technologies (DHTs)“ [28], das App-Entwickler und Entscheidungsträger einen Rahmen gibt, soll sicherstellen, dass neue DHTs klinisch wirksam und ökonomisch sinnvoll sind und gute Evidenzniveaus für digitale Gesundheitsanwendungen gewährleisten. In England gibt es keinen nationalen Erstattungsrahmen für Gesundheitsanwendungen: dies obliegt ausschließlich den CCGs. Zurzeit gibt es keine etablierte Verbindung zwischen der NICE-Überprüfung von digitalen Gesundheitsanwendungen und Erstattungsentscheidungen. Bislang wurden alle NICE-Bewertungen von Gesundheitsapplikationen im Rahmen des Programms MedTech Innovation Briefing (<https://www.nice.org.uk/about/what-we-do/our-programmes/nice-advice/medtech-innovation-briefings>) durchgeführt, das aber keine Empfehlungen abgibt, sondern lediglich die vorhandenen Daten und Erkenntnisse zusammenfasst [24].

- myCOPD ist derzeit die einzige digitale Gesundheitsanwendung, die auf nationaler Ebene explizit über Innovation and Technology Payment (ITP, <https://www.england.nhs.uk/aac/what-we-do/what-innovations-do-we-support/innovation-and-technology-payment/>), eine Initiative zur Innovationsförderung, erstattet wird. myCOPD wurde von über 100 CCGs in England eingekauft.

Mit 2020 begann NICE jedoch mit der systematischen Bewertung von digitalen Gesundheitsanwendungen und startete vier Pilotierungen des „Evidence standards framework“ anhand konkreter Apps:

- *Lifelight First* zur Überwachung von Vitalparametern (<https://www.nice.org.uk/advice/mib213>)
- *myCOPD* für das Selbstmanagement bei chronisch obstruktiven Lungenerkrankungen (<https://www.nice.org.uk/advice/mib214>),
- *Zio XT* zur Erkennung von Herzrhythmusstörungen (<https://www.nice.org.uk/guidance/indevelopment/gid-mt534>) (in Arbeit),
- *DECODE* Entscheidungshilfesystem zur Erkennung von Demenz (<https://www.nice.org.uk/guidance/indevelopment/gid-mt536>) (eingestellt)

NHS Apps Library:
Verzeichnis für sichere
DiGA, aktuell 97 DiGA

individuelle
Rückerstattung in CCGs

Bewertung anhand NHSX
Digital Health Technology
Standards

keine nationale
Rückerstattung

NICE-Bewertungen
ohne Entscheidung
einer Erstattung

myCOPD von vielen
CCGs erstattet

4 Pilotbewertungen
durch NICE

**NICE-Bewertung durch
MTEP-Programm,
ggf. Empfehlung für
Implementierung von DiGA**

Um in den Refundierungsprozess zu gelangen, besteht also zunächst die Notwendigkeit für einen Entwickler von Gesundheitsanwendungen, die App in der NHS Apps Library zu registrieren. Sobald die App registriert ist, kann der Entwickler der Gesundheits-App mit einzelnen Health Trusts und CCGs auf lokaler Ebene verhandeln. Ein nationaler Erstattungsrahmen existiert für Gesundheits-Apps nicht. Wenn eine Gesundheits-App Verbesserungen oder Kosteneinsparungen auf Systemebene nachweisen kann, kann sie vom National Institute of Health and Care Excellence (NICE) im Rahmen seines *Medical Technologies Evaluation Programme* (MTEP, <https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-medical-technologies-evaluation-programme>) evaluiert werden. Eine positive Befürwortung durch das NICE kann (und soll) die Akzeptanz der Gesundheitsanwendung unterstützen. Seit April 2020 richtet der NHS England sogar einen Mechanismus ein, der CCGs dazu auffordert, Technologien zu akzeptieren und zu implementieren, die vom NICE im Rahmen des *Medical Technologies Evaluation Programmes* oder der *Diagnostic Guidance* empfohlen werden. Auch digitale Gesundheitsapplikationen können sich für dieses Programm qualifizieren [24].

3.5.5 Deutschland

**bisher vereinzelt
Selektivverträge**

Bisher werden in Deutschland einige wenige digitale Gesundheitsanwendungen durch einzelne Krankenkassen mittels Selektivverträgen refundiert. Dabei gibt es unterschiedliche Vereinbarungen der Krankenkassen mit den Herstellern [97]:

- Die digitale Gesundheitsanwendung wird den Versicherten der Krankenkasse uneingeschränkt zur Verfügung gestellt,
- Eine Rückerstattung der Kosten der digitalen Gesundheitsanwendung unter bestimmten Konditionen (z. B. nach Indikation) wird vereinbart,
- Ermäßigungen und Teilrückerstattung von digitalen Gesundheitsanwendungen werden vereinbart.

**SkinVision, Preventicus
Heartbeat, Kaia, Tinnitracks,
M-Sense und mySugr**

Beispiele für diese Selektivverträge mit deutschen Krankenkassen sind die digitalen Gesundheitsanwendungen SkinVision, Preventicus Heartbeat, Kaia, Tinnitracks, M-Sense und mySugr, deren Refundierungsmodelle im Detail im Kapitel 3.2 beschrieben sind.

**Digitale-Versorgung-Gesetz
(DVG) und Digitale
Gesundheits-Anwendung-
Verordnung (DiGAV)**

**DiGA-Verzeichnis des
BfArM**

In Deutschland wurde mit dem Digitale-Versorgung-Gesetz (DVG – inkraftgetreten am 19.12.2019) das Verschreiben und die Refundierung von ausgewählten digitalen Gesundheitsanwendungen durch Vertragsärzt*innen und -therapeut*innen beschlossen [11]. Voraussetzung für ein Verschreiben und eine Refundierung einer digitalen Gesundheitsanwendung ist, dass diese im Verzeichnis erstattungsfähiger digitaler Gesundheitsanwendungen (DiGA-Verzeichnis) gelistet ist. Dieses DiGA-Verzeichnis wird durch das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) verwaltet [98]. Anträge zur Aufnahme von digitalen Gesundheitsanwendungen in das Verzeichnis werden durch das BfArM entschieden. Den Antrag für die Aufnahme der digitalen Gesundheitsanwendung in das DiGA-Verzeichnis stellt der Hersteller gegenüber dem BfArM. Die Anforderungen an die digitale Gesundheitsanwendungen und der Ablauf des Aufnahmeverfahrens in das DiGA-Verzeichnis werden durch die Digitale-Gesundheitsanwendungen-Verordnung (DiGAV) des Bundesministeriums für Gesundheit geregelt [98].

Digitale Gesundheitsanwendungen, die in das Verzeichnis aufgenommen wurden können von Vertragsärzt*innen und -psychotherapeut*innen verschrieben werden. Für andere Leistungserbringer, die im Rahmen der GKV Leistungen erbringen (wie Physio- oder Ergotherapeut*innen), ist ein Verschreiben von DiGA nicht vorgesehen [98]. Nach ärztlicher Verordnung können also Nutzer*innen eine digitale Gesundheitsanwendung verschrieben bekommen. Bei Nachweis einer vorliegenden Indikation des Versicherten bei der Krankenkasse können Nutzer*innen eine digitale Gesundheitsanwendung auch ohne ärztliche Verordnung rückerstattet bekommen, solange diese für die entsprechende Indikation im DiGA-Verzeichnis gelistet ist [98].

Für digitale Gesundheitsanwendungen mit geringem Risiko (Medizinprodukt Risikoklasse I oder IIa) wird die Bewertung durch das BfArM durchgeführt, mit der Risikoklasse entsprechenden niedrigeren Evidenzanforderungen. Für digitale Gesundheitsanwendungen mit höherem Risiko (Medizinprodukt Risikoklasse IIb oder III) wird für die Bewertung ein vollständiges HTA durch den G-BA und das IQWiG für die Evidenzbeurteilung durchgeführt (wenn ein Antrag beim G-BA eingereicht wurde), mit der Risikoklasse entsprechenden höheren Evidenzanforderungen.

Nach Eingang des Antrags am BfArM bewertet dieses die digitale Gesundheitsanwendungen innerhalb von 3 Monaten (sog. „Fast-Track-Verfahren“). Dabei werden neben Datenschutz und Benutzerfreundlichkeit insbesondere die Nachweise für positive Versorgungseffekte überprüft [98]:

**Verschreiben durch
Vertragsärzt*innen,
Erbringer*innen von
Psychotherapie**

**Risikoklasse I oder IIa:
Bewertung durch BfArM**

**Risikoklasse IIb oder III:
Bewertung durch G-BA
und IQWiG**

„Fast-Track-Verfahren“

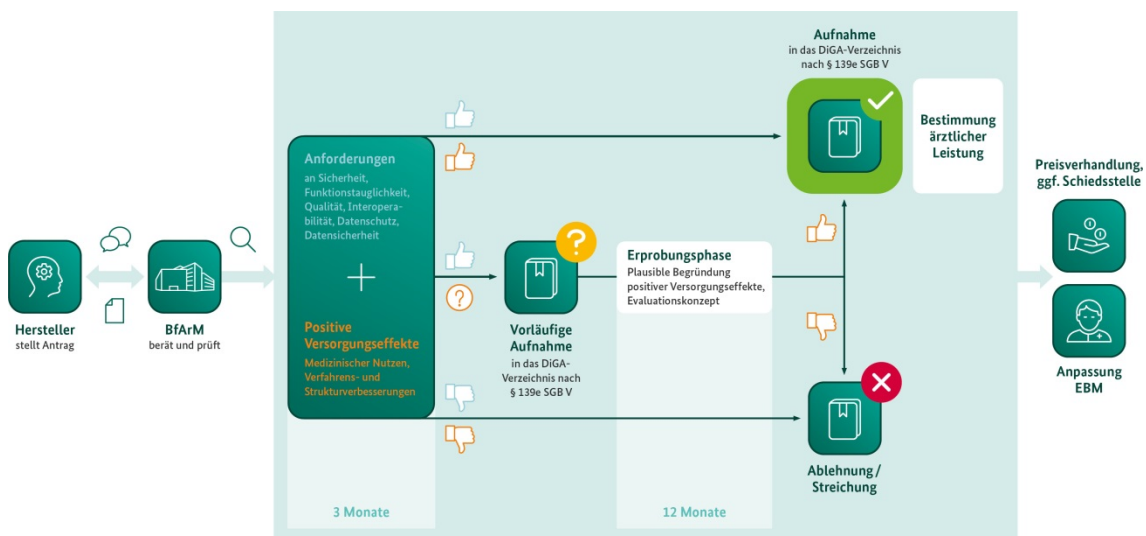


Abbildung 3.5-2: Übersicht zum Ablauf des Fast-Track-Verfahrens [98]

Durch das Fast-Track-Verfahren soll die Integration der digitalen Gesundheitsanwendungen in die Regelversorgung beschleunigt werden: Wenn der Hersteller noch keinen Nachweis für die positiven Versorgungseffekte vorweisen kann, die digitale Gesundheitsanwendung jedoch vielversprechendes Potential aufweist, kann die digitale Gesundheitsanwendung auf Erprobung in das DiGA-Verzeichnis aufgenommen werden [24].

**Aufnahme einer DiGA ins
Verzeichnis auf Erprobung
möglich**

Evidenzgenerierung während Zeit der Erprobung	Dem Antrag auf Erprobung muss ein wissenschaftliches Evaluationskonzept beigelegt werden. Dieses Konzept zur Evaluation muss von einer unabhängigen Institution erstellt werden und allgemein anerkannten wissenschaftlichen Standards entsprechen. Während dieser Zeit der Erprobung muss der Hersteller mit einer Studie einen positiven Versorgungseffekt durch die digitale Gesundheitsanwendung belegen können, die Wahl des Studiendesigns mit dem gewählten Vergleich und die Wahl der Outcomes sind darin zu begründen [98]. Außerdem müssen die Hersteller im Antrag Angaben zu den Themen Datenschutz, Informationssicherheit und Qualität machen, welche über Checklisten der Anlagen 1 und 2 der DiGAV überprüft werden [98].
Studiendesign, gewählter Vergleich und Outcomes relevant	
Positive Versorgungseffekte	Positive Versorgungseffekte können entweder das Vorhandensein eines medizinischen Nutzens oder Struktur- und Verfahrensverbesserungen in der Patient*innenversorgung sein.
medizinischer Nutzen	Der medizinische Nutzen ist im Leitfaden des BfArM folgendermaßen definiert [98]: ■ Verbesserung des Gesundheitszustands ■ Verkürzung der Krankheitsdauer ■ Verlängerung des Überlebens ■ Verbesserung der Lebensqualität Zur Messung des medizinischen Nutzens werden patient*innenrelevante Endpunkte wie Morbidität, Mortalität oder Lebensqualität genannt.
Struktur- und Verfahrensverbesserungen	Die Struktur- und Verfahrensverbesserungen in der Patient*innenversorgung werden im Leitfaden des BfArM folgendermaßen definiert [98]: ■ Verbesserungen bei der Erkennung, Überwachung, Behandlung oder Linderung von Krankheiten, Verletzungen oder Behinderungen, ■ Ausrichtung auf eine Unterstützung des Gesundheitshandelns der Patient*innen oder auf eine Integration der Abläufe der Patient*innen und Leistungserbringern, ■ Verbesserungen in den Bereichen von Behandlungsabläufen, Ausrichtung der Behandlung an Leitlinien, Adhärenz, Zugang zu Versorgung, Patient*innensicherheit, Gesundheitskompetenz, Patient*innensouveränität, Bewältigung krankheitsbedingter Schwierigkeiten im Alltag oder Reduzierung der therapiebedingten Aufwände und Belastungen der Patient*innen und ihrer Angehörigen.
vergleichende Studien empfohlen	Zum Nachweis dieser Struktur- und Verfahrensverbesserungen empfiehlt das BfArM die Durchführung einer vergleichenden Studie an einer über die Indikation definierte Patient*innengruppe, die die digitale Gesundheitsanwendung im Rahmen der Therapie nutzt, im Vergleich zu einer anderen zur selben Indikation definierte Patient*innengruppe die keine digitale Gesundheitsanwendung nutzt [98].
Definition wesentlicher Änderungen der DiGA	Änderungen der digitalen Gesundheitsanwendung, welche wesentlichen Einfluss auf die Bewertungsentscheidung des BfArM haben, müssen vom Hersteller jedenfalls an das BfArM gemeldet werden. Diese wesentlichen Änderungen sind durch das BfArM folgendermaßen definiert [98]: ■ Änderungen, die die Informationen oder Angaben der digitalen Gesundheitsanwendung im Verzeichnis ändern, ■ Änderungen, die die Sicherheit, Funktionstauglichkeit und Qualität der digitalen Gesundheitsanwendung beeinflussen, ■ Änderungen in Bezug auf Datenschutz und Informationssicherheit,

- Änderungen in Bezug auf die positiven Versorgungseffekte (einschließlich Änderungen der Patient*innengruppen für die diese positiven Versorgungseffekte nachgewiesen wurden).

Ab Mai/Juni 2020 war es für Hersteller möglich, Anträge von Herstellern für die Aufnahme einer digitalen Gesundheitsanwendung in das Verzeichnis an das BfArM zu stellen. Die Veröffentlichung des DiGA-Verzeichnisses ist für Herbst 2020 geplant. Im DiGA-Verzeichnis sollen die wesentlichen Informationen zu den einzelnen digitalen Gesundheitsanwendungen für verschreibende Ärzt*innen und Psychotherapeut*innen, aber auch für die Nutzer*innen öffentlich und transparent zugänglich sein.

**Veröffentlichung
DiGA-Verzeichnis
Herbst 2020**

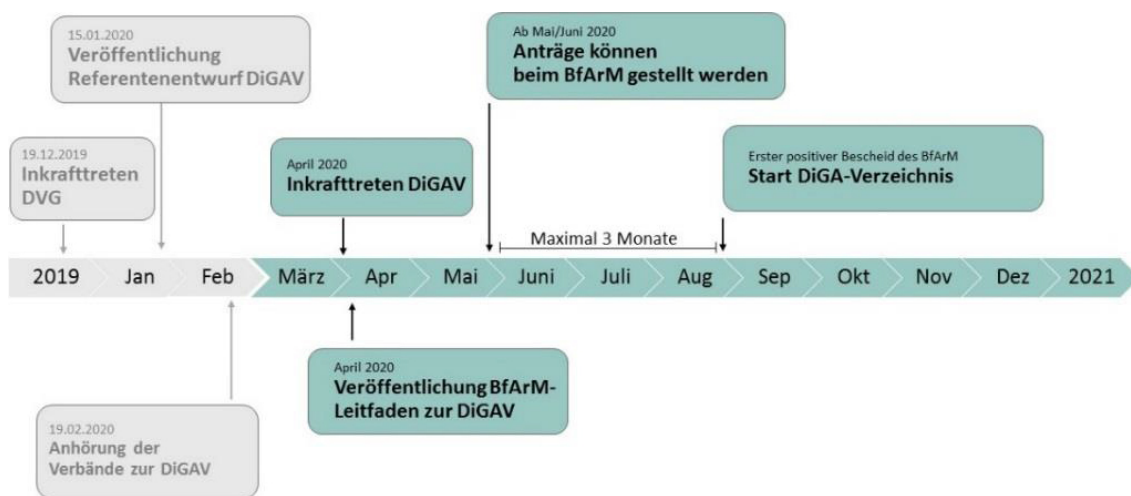


Abbildung 3.5-3: Timeline der Umsetzung des Fast-Track-Verfahrens [98]

Am 06.10.2020 wurden die ersten digitalen Gesundheitsanwendungen in das DiGA-Verzeichnis aufgenommen und veröffentlicht. Die erste App dient der Verhaltenstherapie bei chronischer Tinnitusbelastung („kalmeda“ des Herstellers mynoise GmbH). Die zweite, webbasierte digitale Gesundheitsanwendung dient der Unterstützung von Patient*innen mit bestimmten Angststörungen („velibra“ des Herstellers GAIA AG) [99].

**Erste 2 DiGA in Verzeichnis
aufgenommen**

Weitere digitale Gesundheitsanwendungen sollen in weiterer Folge dem Verzeichnis hinzugefügt werden. Mit den Hersteller*innen von 75 digitalen Gesundheitsanwendungen hat das BfArM bereits Beratungsgespräche durchgeführt. Derzeit befinden sich 21 digitale Gesundheitsanwendungen in der Prüfung durch das BfArM [99].

**Aktuell 21 DiGA in Prüfung
durch das BfArM**

3.5.6 Österreich

In Österreich legte ein Bericht des Institut für Gesundheitsförderung und Prävention (IfGP) 2018 eine Bestandsaufnahme zu Gesundheits-Apps vor [100], in dem zum einen die bislang im deutschsprachigen Raum systematischsten Studien (CHARISMHA [9], Bertelsmann [101], IGES [10]) vorgelegten Typologien für Anwendungen wie auch die Risikoklassen mit erhöhtem Regulationsbedarf beschrieben werden. Darüber hinaus wurde in der genannten Studie eine Befragung von 22 Sozial- und Pensionsversicherungsträgern durch-

**2018:
IfGP-Bericht zu
Gesundheits-Apps**

**16 DiGA in Österreich
angeboten**

geführt (Rücklauf, n=14), die das Ergebnis zeigte, dass bereits 16 digitale Gesundheitsanwendungen in Österreich von den unterschiedlichen Trägerorganisationen angeboten werden (vgl. Tabelle 3.5-1), von denen drei Apps von mehr als einer Organisation angeboten wurden.

Tabelle 3.5-1: In Österreich angebotene Apps (2018) [100]

Nr.	App	Anbietende Organisation(en)	Bei App-spezifischer Erhebung genannt	Bei zusätzlicher Internetrecherche identifiziert
1	Alfa (Anklicken. Loslassen. Freisein. Aufatmen)	VAEB	x	
2	Bildschirmarbeit	AUVA	x	
3	CardioMemory	VAEB, STGKK	x	
4	DiabMemory*	VAEB, STGKK, TGKK	x	
5	EKO2gO	HVB	x	
6	Erste Hilfe Hand	AUVA	x	
7	Food Jungle	HVB	-	x
8	Heben und Tragen	AUVA	x	
9	HerzMobil	STGKK, TGKK	x	
10	Lexikon Prävention	AUVA	x	
11	MAM Elda App	OÖGKK	-	x
12	MedBusters	HVB	x	
13	mySugr Tagebuch-App**	SVA	-	x
14	OÖGKK direkt	OÖGKK	-	x
15	Rauchfrei-App	NÖGKK	x	
16	Schlaganfall Tirol	TGKK	x	
17	Sicherheits- und Gesundheitsmanagement (SGM)	AUVA	x	
18	SOS Tracking.com	BVA	x	
19	Unfallfrei von Anfang an	OÖGKK	-	x
20	...von Anfang an!	OÖGKK	-	x
21	Zahnputz-App	NÖGKK	x	
22	Zeichen der Sicherheit	AUVA	x	
Anzahl angebotener Apps			16	6

* Die Projekte in den Organisationen laufen unter verschiedenen Bezeichnungen (z. B. Gesundheitsdialog Diabetes VAEB, DiabMemory Steiermark, DiabCare Tirol). In der Tabelle wird der Name der entsprechenden App angeführt.

** Die SVA übernimmt für ihre Versicherten bei Teilnahme am Disease-Management Programm „Therapie Aktiv“ die Kosten für die mySugr Tagebuch-App und die mobile und webbasierte mySugr Typ-2 Diabeteschulung (SVA, o.Ź.; mySugrGmbH, o.Ź.).

DiGA zur Gesundheitsförderung und Prävention, aber auch therapeutische Funktionen und Datenverarbeitung

Wie die Entscheidungen zugunsten der Apps gefällt wurden und ob die Fusionierung der Sozialversicherungen bislang einen Einfluss auf das Spektrum der angebotenen Apps hatte, ist nicht bekannt, resp. wurde für diesen Bericht nicht recherchiert. Einige der digitalen Gesundheitsanwendungen wurden von den Trägern auch mitfinanziert. Die Mehrheit der angebotenen Apps sind zur Gesundheitsförderung und Prävention und Serviceanwendungen (also niedrige Risikoklassen), während fünf Apps auch therapeutische Funktionen übernehmen (u. a. mySugr und DiabMemory), Daten erfassen, individuelle Verläufe darstellen und mit dem/der behandelnden Arzt/Ärztin kommunizieren (Feedback einholen), drei Apps verarbeiten die Daten auch

[100]. 15 der 22 digitalen Gesundheitsanwendungen werden den Versicherten kostenfrei zur Verfügung gestellt, fünf weitere einem nur eingeschränkten Personenkreis (chronisch Kranke unter den Versicherten) oder mit Zukauf von Zusatzfunktionen.

Ein vom Hauptverband der Sozialversicherungen 2018 in Auftrag gegebener Review befasste sich mit dem Kenntnisstand in der ökonomischen Evaluation von eHealth Applikationen [102], um erste Empfehlungen für die ökonomische eHealth-Evaluation in Österreich zu erarbeiten. Eine Entwicklung formeller Entscheidungsstrukturen für die Identifikation, Priorisierung, Implementierung und Skalierung von eHealth unter Einbezug relevanter Stakeholder auf allen politischen Ebenen in Österreich wird empfohlen [102].

Derzeit ist ein „digitaler Gesundheitspfad“ [103] in Planung, der die Entwicklung einer integrierten Plattform für unterschiedliche eHealth und mHealth Anwendungen und deren Koordination zur Aufgabe hat und einen zertifizierten und standardisierten Zugang ins österreichische Gesundheitswesen ermöglichen soll. Darüber hinaus nimmt der HVB, jetzt Dachverband an Arbeitsgruppen der European Social Insurance Platform (ESIP) teil, in der auch Bewertungsschemata für mHealth Applikationen diskutiert werden.

Ein ausgearbeitetes Konzept für ein standardisiertes Vorgehen liegt bislang noch nicht vor.

**2018:
Review zur ökonomischen
Evaluation von eHealth**

**„digitaler
Gesundheitspfad“,
ESIP-Arbeitsgruppen**

**ausgearbeitetes Konzept
noch ausständig**

4 Diskussion

4.1 Zusammenfassung: Kritische Bewertung der Frameworks und ihrer Anwendbarkeit

Für diesen Projektbericht wurden exemplarisch sechs Assessment Frameworks für digitale Gesundheitsanwendungen, die ab 2018 publiziert wurden, ausgewählt und auf Ähnlichkeiten und Unterschiede analysiert und in ihrer Anwendbarkeit getestet. Die Auswahl der Assessment Frameworks erfolgte nach deren Aktualität, Reifegrad und voraussichtlicher Anwendbarkeit. Assessment Frameworks, die einen zu engen Scope (z. B. nur Evaluation der Benutzerfreundlichkeit, oder Beschränkung auf einzelne Indikationen) oder keine ausreichende Aktualität in Bezug auf mHealth hatten, wurden nicht gewählt.

Die internationalen Assessment Frameworks bedienen sich sehr unterschiedlicher Begriffe für ihren Anwendungsbereich („Mobile Medical Applications“, „Digital Health Technologies“, „Digital Health Products“). Die beschriebenen Frameworks basieren zum Teil auf systematischen Literaturanalysen zu Bewertungskriterien und bestehenden Frameworks zur Bewertung von nicht-digitalen Technologien (HTA-Domänen), sowie bestehenden Technologie-Entwicklungsmodellen (MOST, IDEAS) und wurden – zum Teil – von verschiedenen interdisziplinären Interessensgruppen konzipiert und diskutiert. Die beschriebenen Assessment Frameworks nutzen unterschiedliche Formate (Checklisten, Fragebögen, Tabellen) und weisen Unterschiede hinsichtlich des Spektrums der Bewertungsdomänen, zusätzlicher technologiespezifischer Aspekte und Zielgruppen auf. Keines der Frameworks ist vollständig etabliert oder ist Praxis-erprobt.

Von den beschriebenen Assessment Frameworks schlägt nur das *Evidence Standards Framework* von NICE [28] eine Kategorisierung nach definierten Risikoklassen der digitalen Gesundheitsanwendungen vor. Diese Einteilung erscheint unerlässlich, um aus der Vielzahl unterschiedlicher digitaler Gesundheitsanwendungen Entscheidungen nach Art und Umfang von Evidenz zu treffen. Die anderen in diesem Projektbericht beschriebenen Assessment Frameworks benennen zwar das Risiko möglicher gravierender klinischer Konsequenzen durch mHealth Anwendungen, schlagen jedoch keine Einteilung nach definierten Risikoklassen vor.

In einem Abgleich der derzeitigen Risikoklassen-Einstufung gemäß der alten CE-Kennzeichnung und den Risikoklassen nach der neuen MDR/IVDR [18, 19] wird offensichtlich, dass die Mehrheit der hier ausgewählten digitalen Gesundheitsanwendungen nach der neuen Verordnung in einer höheren Risikoklasse wären. Auch die Überprüfung der bislang vorliegenden Nutzenbelege in Form wissenschaftlicher Studien ergab, dass fast ausschließlich retrospektive Anwendungsbeobachtungen und Datenanalysen von Real-World Daten, bestenfalls nach einem prospektiven Studienplan – in den seltensten Fällen mit einer Kontrollgruppe (gegenüber SOC/standard-of-care) – durchgeführt wurden. Jüngere, noch laufende Studien weisen dagegen bereits höherwertige Studiendesigns vor. Bei digitalen Anwendungen mit Anspruch/Erwartung an gleiche diagnostische Genauigkeit wie Fachärzt*innen wären für den Nachweis der Spezifität und Sensitivität der Diagnostik eines Algorithmus saubere, prospektive Diagnostikstudien wünschenswert und es ist zu erwarten, dass diese auch unter der neuen Medizinprodukteverordnung bei höheren Risikoklassen häufiger werden.

**Detailanalyse von
6 ausgewählten
Assessment Frameworks**

**große Heterogenität der
Assessment Frameworks**

**NICE Framework:
Einteilung in Risikoklassen**

**Höherstufung des Risikos
der DiGA nach der neuen
MDR/IVDR**

Deutschland und England in der Regulierung weit fortgeschritten	In einer Länderanalyse zeigt sich, dass Deutschland der gegebenen Situation (unzählige verfügbare Apps, geringe Evidenz zu Nutzenbelege) Rechnung trägt und nicht nur eine Registrierung möglicher verschreibbarer mHealth Applikationen regelt, sondern darüber hinaus die Apps nach Risikoklassen – orientiert an der Medizinprodukteverordnung – kategorisiert und für höhere Risikoklassen (IIb und III) entsprechende Nutzenbelege durch Studien vorsieht. Diese können in Erprobungsstudien – nach Zulassung – erbracht werden. Deutschland ist mit dieser sehr klaren Vorgangsweise und der Einräumung der Möglichkeit prospektiver Evaluationsstudien neben England in Europa – soweit den Autor*innen bekannt – in der Regulierung am weitesten fortgeschritten.
Risikoklassifizierung bereits durch IGES Studie vorgeschlagen	Die 6-teilige Studie der Bertelsmann-Stiftung (2016-2019) [101], die CHARISMHA Studie (2016) [9] und die IGES Studie (2016) [10] legten für diese klaren Regelungen das wissenschaftlich-analytische Fundament. Diese drei großen Studien befassten sich im Detail mit der Herangehensweise zur Bewertung von digitalen Gesundheitsanwendungen. Bereits 2016 (also lange vor der Veröffentlichung des NICE Frameworks) wurde im IGES Bericht, der im Auftrag der Techniker Krankenkasse erstellt wurde, vorgeschlagen, als ersten Schritt eine Risikoklassifizierung vorzunehmen. Anwendungen, die einen Gesundheitsdienstleister ersetzen, sind immer in Risikoklasse 3 einzustufen [10]. Als Refundierung wird – nach Vorlage eines Nutzenbelegs – eine kollektive (für alle Versicherten) oder eine selektive (für bestimmte Gruppen, wie etwa chronisch Kranke) Abgeltung vorgeschlagen. Darüber hinaus könnten – sofern Nutzenbelege noch ausstehen – selektive Verträge abgeschlossen werden, die unter Studienbedingungen Datenerhebungen etwa zu bestimmten Nutzer*innengruppen durchführen und Evidenz generieren. Wie im Länderbericht zu Deutschland ersichtlich, werden diese Vorschläge weitgehend umgesetzt. Ein „schleichendes Akzeptieren“ von eHealth Interventionen wird hingegen abgelehnt: als negatives Beispiel wird das App zur Behandlung von Tinnitus genannt [104].
kollektive oder selektive Abgeltung	
begleitende Generierung von Evidenz	
Umgang mit gleichwertigen Apps noch ungeklärt	Der Umgang mit gleichwertigen Apps, auch als Me-too oder „Trittbrettfahrerproblematik“ bekannt, ist noch ungeklärt und lässt sich nur „politisch klären“ [104], indem – ähnlich dem Patentschutz – dem Hersteller ein befristeter Marktvorteil gewährt wird oder jedem Hersteller, der eine öffentliche Refundierung anstrebt, eine entsprechende Nutzenbewertung abverlangt wird [104].
Regelung zur Haftung noch ausständig	Auch eine Regelung zur Haftung, die über die reine Produkthaftung (fehlerhafte Produkttechnik) hinausgeht, also eine Haftung für eine fehlerhafte Diagnose, wie etwa das Übersehen einer Erkrankung mit schwerwiegenden Folgen, steht noch aus [104].
kostenlose DiGA ohne Regulation	Ein weiteres Problem stellen digitale Gesundheitsanwendung dar, die vom Hersteller kostenlos (etwa als Beigabe der Standardsoftware des Betriebssystems des mobilen Gerätes) zur Verfügung gestellt werden, aber Schaden anrichten könnten. Diese Art von digitalen Gesundheitsanwendungen könnte man über die Refundierungshürde nicht aus dem System fernhalten.
mHealth Assessment erfordert Zusammenarbeit mehrerer Organisationen	Eine 2018 publizierte Stellungnahme der Weltorganisation von Herstellern digitaler mobil-Technologien (https://mobileworldcapital.com/) zum Assessment von mHealth betont die Notwendigkeit der Zusammenarbeit von Entwickler*innen der Technologien mit den Evaluators*innen. Laut den Autor*innen bedarf es eines globalen Assessment Frameworks mit einer klaren Spezialisierung auf Teilbereiche, da eine einzige Organisation die Vielfalt der Aufgaben eines mHealth Assessments nicht bewältigen könne [105].

4.2 Weitere Aspekte und Bereiche mit besonderer Aufmerksamkeit

4.2.1 Datenschutz

Im Vergleich zu anderen Medizinprodukten spielt der Datenschutz bei digitalen Gesundheitsanwendungen eine noch bedeutendere Rolle. Durch die digitalen Technologien ist eine kontinuierliche Erhebung von Messwerten sowie Tracking von Symptomen möglich. Darüber hinaus ist eine Zuordnung zu einem personenbezogenen Datenprofil und eine Weiterverarbeitung und Speicherung der Daten ohne zeitliche Einschränkung möglich.

Der persönliche Zugang zum Thema Datenschutz in digitalen Gesundheitsanwendungen variiert je nach Nutzer*in. So könnte für eine Nutzer*in die Aufzeichnung, die Verarbeitung und der Verkauf von Standort- und Bewegungsdaten keine Rolle spielen, während für eine andere Nutzer*in eine solche Verwendung von Daten nicht in Frage käme [45].

Problematisch ist, dass manche Hersteller*innen von digitalen Gesundheitsanwendungen die Daten der Nutzer*innen ohne deren Kenntnis weiterverkaufen [27]. Gesundheitsbezogene Daten können durchaus Auswirkungen auf Versicherungstarife, Kreditwürdigkeiten oder Jobchancen haben.

In einem 2014 publizierten Green Paper der europäischen Kommission wird das Misstrauen der Nutzer*innen gegenüber digitalen Gesundheitsanwendungen mit Bedenken in Bezug auf fehlenden Datenschutz begründet [106]. Dieses Misstrauen scheint nicht unbegründet: Laut einem Positionspapier des deutschen Netzwerks für Versorgungsforschung ist davon auszugehen, dass kostenlose digitale Gesundheitsanwendungen in der Regel Daten der Nutzer*innen für wirtschaftliche Zwecke verwenden [17].

Eine Zusammenarbeit zwischen der Techniker Krankenkasse und der digitalen Gesundheitsanwendung Ada Health wurde aufgrund von Verdacht bezüglich Datenschutzmängel beendet. Dabei soll das Unternehmen Ada Health die personenbezogenen Gesundheitsdaten der Nutzer*innen an Facebook und den US-Dienstleister Amplitude übertragen haben. Das Unternehmen Ada Health weist diese Vorwürfe zurück und beschreibt, dass die Übergabe der Daten notwendig sei um die Patient*innensicherheit zu gewährleisten und um Verbesserungen an der digitalen Gesundheitsanwendung durchführen zu können [107].

Ein Beispiel für Probleme mit der Datensicherheit ist die digitale Gesundheitsanwendung Babylon Health. Diese digitale Gesundheitsanwendung bietet den Nutzer*innen ein Überprüfen der Symptome bei anschließenden telemedizinischen Video-Konsultationen mit Therapeut*innen. Diese Video-Konsultationen werden innerhalb der digitalen Gesundheitsanwendung gespeichert, um von den Nutzer*innen bei Bedarf erneut abgespielt werden zu können. Durch ein Datenleck im Juni 2020 waren für Nutzer*innen von Babylon Health die Video-Konsultationen mit sensiblen Inhalten anderer Nutzer*innen sichtbar [108].

An dieser Stelle soll erwähnt werden, dass die Aspekte Zugriffsrechte und Datensicherheit im NICE Evidence Standards Framework für eHealth Technologien der Gruppe 1 noch nicht in Betracht gezogen werden, sondern eine Bewertung dieser Aspekte erst für eHealth Technologien der Gruppe 2 erfolgt [29].

**Datenschutz bei DiGA
besonders relevant**

**Haltung zu Datenschutz
individuell**

**Verkauf von
Gesundheitsdaten**

**Misstrauen der
Nutzer*innen bei
kostenlosen DiGA**

**Datenschutzmängel
beendeten
Zusammenarbeit**

**aufgezeichnete
Video-Konsultationen von
Datenleck betroffen**

**NICE Framework:
Datenschutz nicht für alle
Evidenzstufen bewertet**

Interoperabilität: Datenschutz und Datensicherheit	Die meisten digitalen Gesundheitsanwendungen sind nicht untereinander verknüpft [88]. Dabei können digitale Gesundheitsanwendungen erst bei ausreichender Interoperabilität ihr volles Potenzial entfalten [109]. Allerdings birgt eine erhöhte Interoperabilität auch Risiken in Bezug auf Datenschutz. Analog zu höheren Medizinprodukt-Risikoklassen, welche mehr Evidenz zu klinischem Nutzen erfordern, bedarf eine höhere Interoperabilität ein höheres Maß an Datenschutz und Datensicherheit [110] und müssten daher bereits im Assessment Berücksichtigung finden.
Änderungen im Datenschutz können Neubewertung veranlassen	Fraglich bleibt, wie genau mit digitalen Gesundheitsanwendungen umgegangen werden soll, die im Verlauf keinen ausreichenden Datenschutz oder Probleme in Bezug auf Datensicherheit aufweisen. Ein Ausschluss oder Aussetzen der Refundierung könnte die Hersteller dazu motivieren, ausreichend Aufmerksamkeit und Ressourcen auf diese Aspekte zu wenden. Im DiGA-Leitfaden des BfArM [98] werden Änderungen in der digitalen Gesundheitsanwendung in Bezug auf Datenschutz und Informationssicherheit als wesentliche Änderungen genannt, welche dem BfArM zu melden sind und eine Neubewertung veranlassen können.
4.2.2 Künstliche Intelligenz	
diagnostische Fertigkeit von K.I. rasant entwickelt	Die Fähigkeiten von künstlicher Intelligenz entwickeln sich rasant. Durch den Einsatz von Deep-Learning-Algorithmen ist die diagnostische Fertigkeit in manchen Teilbereichen der Medizin auf dem Niveau von Fachkräften. So zeigte eine systematische Übersichtsarbeit und Meta-Analyse von Liu et al., dass Deep-Learning-Algorithmen in der medizinischen Bildauswertung ähnliche Ergebnisse zu Fachkräften erzielten, wobei die Studien laut den Autor*innen methodische Schwächen aufweisen und standardisiertere Studien notwendig seien [111].
	Es ist zu erwarten, dass sich die Fähigkeiten von künstlicher Intelligenz zunehmend weiter verbessern und in weiteren medizinischen Teilbereichen ausreichende Ergebnisse erzielen werden.
Ethikrat: Algorithmen sollen transparent und nachvollziehbar sein	Der Deutsche Ethikrat empfiehlt den Einsatz von Algorithmen transparent offenzulegen und für die Nutzer*innen der Technologie nachvollziehbar zu erläutern. Diese Erklärung sollte folgende Aspekte beinhalten [112]; ■ Welche Nutzerdaten erhoben werden und für welche Analysen, Entscheidungs- oder Auswahlprozesse einfließen; ■ Welche Nutzerdaten ausdrücklich nicht erhoben werden (z. B. zur Vermeidung von Diskriminierung); ■ Welche Ableitungen, Vorhersagen oder Entscheidung auf Basis dieser Daten mittels Algorithmen getroffen werden; ■ Ob ein Datenprofil mit abgeleiteten Größen der Nutzer*innen erstellt wird und dessen erwartete Aussagekraft; ■ Ob und in welcher Form anonymisierte personenbezogene Daten in Modelle einfließen, wer über Nutzungsrechte dieser anonymisierten Daten verfügt.
„CONSORT-AI“ Statement Erweiterung für Interventionen mit K.I.	Im September 2020 wurde eine Erweiterung des CONSORT Statements (Leitlinie zur verbesserten Darstellung von randomisierten Studien, bestehend aus 25 Punkten) veröffentlicht. In dieser Erweiterung werden relevante Informationen zur Anwendung von Artificial Intelligence in klinischen Studien abgefragt. Die sogenannte CONSORT-AI Erweiterung (Consolidated Standards

of Reporting Trials-Artificial Intelligence) soll die Planung und die Berichterstattung von klinischen Studien über Interventionen die künstliche Intelligenz verwenden verbessern und dadurch die Patient*innensicherheit erhöhen [113].

Für die meisten in diesem Bericht beschriebenen digitalen Gesundheitsanwendungen konnte keine eindeutige Aussage zur Nutzung künstlicher Intelligenz getroffen werden. Einige der Hersteller geben zwar an, dass die Funktionen der digitalen Gesundheitsanwendung auf künstlicher Intelligenz und Machine learning Algorithmen basieren, genaue Angaben zur Funktionsweise oder transparente Erklärungen zu den verwendeten Algorithmen werden jedoch nicht öffentlich gemacht. Dadurch war eine Zuordnung der digitalen Gesundheitsanwendungen zu fixierten oder adaptiven Algorithmen nur nach bedingter Einschätzung möglich (vgl. Tabelle 3.2-1).

DiGA:
Algorithmen nicht
transparent beschrieben

4.2.3 Digital Health Literacy

Die Digital Health Literacy spielt bereits bei der Auswahl einer geeigneten digitalen Gesundheitsanwendung eine Rolle. Nicht nur für Nutzer*innen, sondern auch für Therapeut*innen ist es dabei eine große Herausforderung aus der Vielzahl von digitalen Gesundheitsanwendungen eine geeignete und sichere Anwendung zu wählen. Derzeit wählen viele Nutzer*Innen digitale Gesundheitsanwendungen allein anhand der Sternbewertungen in den App Stores aus, diese haben jedoch nur eine eingeschränkte Aussagekraft [45].

digitale
Gesundheitskompetenz
bei Auswahl von DiGA

Zur Steigerung der Datensouveränität empfiehlt der Deutsche Ethikrat das Fördern von digitaler Bildung. Diese Nutzer*innenkompetenz sollte schon in der Schule vermittelt werden, da Kinder bereits digitale Anwendungen nutzen und dabei Daten generieren. Auch in Unternehmen und Institutionen sind regelmäßige entsprechende Schulungen empfehlenswert [112]. Es existieren bereits Checklisten die sich an Nutzer*innen richten, um die Vertrauenswürdigkeit von digitalen Gesundheitsanwendungen besser einschätzen zu können. Beispiele für solche Checklisten sind die „mHealth App Trustworthiness Checklist“ oder die „HealthOn Checkliste für Gesundheits-Apps“ [114, 115].

digitale Bildung,
Checklisten für
Nutzer*innen

4.2.4 Neue Studiendesigns

Für die in dieser Arbeit untersuchten digitalen Gesundheitsanwendungen fanden sich sogenannte real-world-evidence (RWE) Datenbankanalysen. Bei dieser Art von Studie werden Daten innerhalb der digitalen Gesundheitsanwendung vom Anbieter erhoben und im Sinn einer Beobachtungsstudie analysiert ohne aber die zahlreichen Counfounder und Verzerrungsrisiken zu berücksichtigen.

RWE-Datenbankanalysen

Verschiedene Autor*innen haben begonnen, über die methodischen Herausforderungen bei der Überprüfung der Wirkungen von digitalen Gesundheitsanwendungen zu reflektieren. Ein systematischer Review zu 27 Studien zu mHealth bei psychiatrischen Patient*innen zeigte hohe Heterogenität in Studiendesigns, Patientenrekrutierung, Wahl der Komparatoren und Endpunkte sowie und niedrige Qualität der Evidenz [116]. Ein weiterer Review befasste sich mit der Verteilung der Designs von mHealth – zumeist akademischen – Studien: entgegen den Erwartungen der Autor*innen war die Mehrheit (80 %

hohe Heterogenität
in Studiendesigns

kleine Studien,
kurze Dauer

von 71 identifizierten Studien) RCTs, sowie Pre- und Postinterventionsbeobachtungen, aber auch rein qualitative Studien [117]. Die meisten Studien hatten eine kleine Teilnehmer*innenzahl (ca 112) und eine kurze Dauer.

**„micro-randomized trials“
(MRTs)**

**„sequential multiple
assignment randomized
trials“ (SMARTs)**

Zunehmend werden neue experimentelle Studiendesigns, die die kurzen Innovationszyklen der mHealth Applikationen berücksichtigen, entwickelt: es werden etwa „micro-randomized trials (MRT)“ vorgeschlagen [118-121], die adaptive und Kontext-abhängige („just-in-time“) Interventionen mit einer Modellierung von Kausalitäten von Effekten kombinieren wollen. Als weiteres Studiendesign kommt „Sequential Multiple Assignment Randomized Trial (SMART)“ [122] zum Vorschlag, bei dem individuelle Variablen (etwa Schweregrad, Präferenzen) und individuelle Endpunkte (wie Therapieantwort, Adhärenz) in die Bewertung einfließen.

Methodische Weiterentwicklungen in diesem Gebiet komplexer mHealth Anwendungen konkurrieren mit anderen Interventionen und sollten daher dieselben Evidenzanforderungen erfüllen (müssen), insbesondere da große Erwartungen (Entlastung bei Spitalsaufenthalten von chronisch Kranken durch Monitoring von Vitalwerten, hohe diagnostische Genauigkeit von Algorithmen) an sie gestellt werden, die es zu bestätigen (oder verwerfen) gilt [123]. Nicht auszuschließen sind negative Auswirkungen wie Schaden durch fehlerhafte Diagnostik (Screening-Apps), die unnötigen Ressourcenverbrauch durch Abklärungsdiagnostik zur Folge haben.

4.3 Limitationen der Arbeit

**keine stringente
systematische
Übersichtsarbeit**

Eine Limitation der Arbeit ist, dass diese nicht stringent der Methodik einer systematischen Übersichtsarbeit entspricht. Die Suche nach internationalen Bewertungsinstrumenten erfolgte zwar durch eine systematische Literatursuche, die Auswahl der analysierten Bewertungsinstrumenten wurde jedoch exemplarisch nach Aktualität, Ausgereiftheit und zu erwartender Anwendbarkeit der Bewertungsinstrumente getroffen.

**keine Daten zur
Verwendung der
Frameworks**

Eine weitere Limitation der Arbeit liegt darin, dass nicht untersucht wurde, inwieweit die beschriebenen Assessment Frameworks für die Evaluation von digitalen Gesundheitsanwendungen tatsächlich bereits verwendet werden, und inwiefern die Assessment Frameworks für Refundierungsentscheidungen herangezogen werden.

**Auswahl der digitalen
Gesundheitsanwendungen
nicht repräsentativ**

Eine weitere Limitation liegt in der Auswahl der digitalen Gesundheitsanwendungen. Aus der Vielzahl der verfügbaren digitalen Gesundheitsanwendungen wurde eine Auswahl für die Analyse getroffen, die Ergebnisse dieser Analyse sind daher als exemplarisch (und nicht repräsentativ) einzuordnen. An dieser Stelle soll nochmal betont werden, dass das Gebiet der digitalen Gesundheitsanwendungen und der verfügbaren Evidenz ein äußerst dynamisches ist und die in diesem Bericht beschriebene Datenlage eine Momentaufnahme ist.

**Länderbeschreibung
abhängig von verfügbaren
Informationen**

Die Auswahl der Länder für die Beschreibung der unterschiedlichen Herangehensweisen, Strategien und Refundierungsmodelle für digitale Gesundheitsanwendungen war von der Verfügbarkeit der Informationen geprägt.

Aufgrund dieser Limitationen sind Aussagen zur Anwendbarkeit der Assessment Frameworks eingeschränkt. In der Gegenüberstellung der vorliegenden Evidenz digitaler Gesundheitsanwendungen mit den Evidenzerfordernissen der jeweiligen Assessment Frameworks erachten wir keine gravierenden Verzerrungen in der Darstellung.

4.4 Schlussfolgerungen

Im Folgenden fassen wir nochmals unsere Schlussfolgerungen zusammen, die letztendlich zu einer Empfehlung führen:

- a. Es existiert ein großes Angebot an digitalen Gesundheitsanwendungen, welche aufgrund von Fortschritten auf dem Gebiet der künstlichen Intelligenz neben rein informativen zunehmend diagnostische und therapeutische Funktionen anbieten. Für die Mehrzahl der hier untersuchten digitalen Gesundheitsanwendungen liegt wenig Evidenz zum tatsächlichen Nutzen vor. Mit der CE-Zertifizierung nach MDR/IVDR und der entsprechenden Risikoklasseneinstufung sollte auch bessere Evidenz vorgelegt werden. Jedenfalls ist die Mehrheit der hier ausgewählten digitalen Gesundheitsanwendung nach der „neuen“ MDR 2017/745 in eine höhere Risikoklasse (min 2a) einzuteilen als unter der „alten“ MDD-Zulassung.
- b. Die Medizinprodukteverordnung (MDR/IVDR) wird aber auch in Zukunft „nur“ das Nutzen-Schaden-Verhältnis für die Zulassung beurteilen, während HTA die vergleichende Nutzen-Schaden-Relation im Kontext bereits etablierter Technologien und Interventionen ins Zentrum stellt und dementsprechend nicht nur der medizinische Nutzen, sondern auch die Kosteneffektivität, organisatorische und soziale Auswirkungen von mHealth Interventionen Bedeutung haben. Vergleiche mit dem bisher existierenden Versorgungsstandard sind in Form von kontrollierten Studien nicht immer, aber in rezenten Studien (z. B. laufende RCT-Studien) häufiger gegeben.
- c. Die meisten Assessment Frameworks sehen keine Einteilung in definierte Risikoklassen der digitalen Gesundheitsanwendung vor. Das Evidence Standards Framework von NICE bietet eine genaue Einteilung nach definierten Risikoklassen, welche für die Vielzahl an digitalen Gesundheitsanwendungen notwendig erscheint. Höhere Risikoklassen erfordern demnach ein höheres Maß an Evidenz. Das NICE Framework wird von den meisten HTA-Institutionen als „best practice“ erachtet.
- d. Einzelne digitale Gesundheitsanwendungen werden von verschiedenen regionalen Krankenkassen bereits refundiert, darüber hinaus ist eine Entwicklung zu beobachten, Regelungen auf nationaler Ebene zu definieren und institutionelle Zuständigkeiten gesetzlich zu verankern. Dadurch ergeben sich Möglichkeiten für eine stärkere Regulierung und für eine gezieltere Prüfung von Nutzen und Sicherheit der einzelnen Anwendungen. Die gründliche Evaluation von digitalen Gesundheitsanwendungen (unter anderem der strengen Datenschutzvorgaben) erfordert die Zusammenarbeit verschiedener Institutionen mit entsprechender Aufgabenverteilung.

großes Anbot an DiGa mit auch diagnostischen & therapeutischen Funktionen

**Unterschied MDD vs. MDR/IVDR
CE Zertifizierung sollte bessere Evidenz liefern**

Zulassung vs. HTA

**vergleichende Nutzen-Schaden Bewertung
+ Kosteneffektivität
+ organisatorische und soziale Auswirkungen**

zunehmend häufiger mHealth in RCTs

NICE Framework stellt DiGA Risikoklassen Evidenzanforderungen gegenüber = best practice

eine grundsätzliche Regelung auf nationaler Ebene geht den Refundierungen voraus

gesetzliche Verankerung, Aufgabenverteilung

künstliche Intelligenz
hat noch Bedarf nach
Regelung:

Transparenz

digital literacy

- e. Aspekte der künstlichen Intelligenz werden in den Assessment Frameworks nicht ausreichend thematisiert. Es besteht aber Notwendigkeit, durch Auflagen zur Offenlegung der Grundlage der Algorithmen (detaillierte Technologiebeschreibung der AI-Algorithmen) die Hersteller zu mehr Transparenz zu verpflichten. Die Berücksichtigung der vom Deutschen Ethikrat definierten Kriterien im Umgang mit künstlicher Intelligenz sowie die Anwendung des CONSORT-AI Statements bei RCTs können zu Verbesserungen der Transparenz der Berichterstattung betragen.
- f. Nicht zuletzt ergibt sich ein Mehrbedarf an Strategien zur Steigerung der digitalen Gesundheitskompetenz der Anwender*innen.

5 Empfehlung für Framework für Österreich

In den Analysen zu den international verfügbaren Assessment Frameworks zu mHealth und deren exemplarische Anwendung an den für diesen Bericht ausgewählten digitalen Gesundheitsanwendungen zeigte sich, dass jedes Assessment Framework in unterschiedlichem Ausmaß Limitationen in der Anwendung aufweist. Daraus folgt, dass für eine Evaluation von digitalen Gesundheitsanwendungen die Kombination mehrerer Assessment Frameworks vorteilhaft ist.

Es empfiehlt sich ein abgestuftes Vorgehen. Dabei sollte zunächst eine Einteilung nach Risikoklasse der digitalen Anwendung erfolgen, um anschließend die relevanten HTA-Domänen zu bestimmen und diese zur Bewertung heranzuziehen. Dieses Vorgehen scheint zeit-effizient. Folgender Prozess wird vorgeschlagen:

1. Überprüfung, ob eine CE-Kennzeichnung vorliegt und nach welcher Risikoklasse (Medizinprodukteverordnung EU MDR 2017/745, insb. Regel 11) die digitale Gesundheitsanwendung klassifiziert wurde.
2. Abgleich und Verifizierung der MDR-Einteilung mit den Risikoklassen nach NICE (vgl. Tabelle 3.4-1). Zu den Evidenzerfordernissen nach jeweiliger Risikoklasse empfehlen wir die Anwendung des *NICE Evidence Standards Frameworks* (vgl. Anhang 7.1.2).
3. Screening nach und Scoping der vorhandenen Evidenz und Abgleich der Studiendesigns mit den Erfordernissen an Studiendesigns der jeweiligen Risikoklasse. Mit dem derzeitigen Versorgungsstandard vergleichende (kontrollierte) Studien stellen das höchste Maß an Evidenz dar; retrospektive Auswertungen von Nutzer- und Nutzungsdaten können hingegen nur Hypothesen-bildend sein.
4. Evaluation unter Berücksichtigung relevanter HTA-Aspekte, die sich an den Erwartungen zum Potential und Befürchtungen zu Folgewirkungen orientieren. Zur Berücksichtigung der relevanten HTA-Aspekte empfehlen wir die Anwendung des *HTA Module for Mobile Medical Applications* (vgl. Anhang 7.1.1) sowie die Anwendung des *Digi-HTA Frameworks* (vgl. Anhang 7.1.3). Zur Bewertung von Studien, die medizinische Interventionen mit künstlicher Intelligenz überprüfen, empfehlen wir die Anwendung der CONSORT-AI Erweiterung (vgl. Kapitel 4.2.2) als Orientierungshilfe.
5. Wir empfehlen, diesen Prozess zu pilotieren und ggf. zu adaptieren.

**Kombination mehrerer
Assessment Frameworks
vorteilhaft**

abgestuftes Vorgehen:

**Überprüfung der
CE-Kennzeichnung und der
Risikoeinstufung nach MDR
Abgleich mit NICE Evidence
Standards Framework:
Risikoklassen und Evidenz**

**Screening/Scoping nach
vorliegender Evidenz**

**Evaluation nach
relevanten Aspekten:**

**HTA Module for Mobile
Medical Applications
Digi-HTA Framework**

Pilotierung des Prozesses

6 Literatur

- [1] Our World in Data. Daily hours spent with digital media, United States 2008-2018. 2020 [cited 04.10.2020]. Available from: <https://ourworldindata.org/grapher/daily-hours-spent-with-digital-media-per-adult-user>.
- [2] Berg A. Bitkom Research – Digital Health. 2019 [cited 04.10.2020]. Available from: https://www.bitkom.org/sites/default/files/2019-05/190508_bitkom-pressekonzferenz_e-health_prasentation.pdf.
- [3] Rohleder B. Bitkom Research – Digital Health. 2020 [cited 04.10.2020]. Available from: https://www.bitkom.org/sites/default/files/2020-07/prasentation_digitalhealth2020.pdf.
- [4] Oh H., Rizo C., Enkin M. and Jadad A. What is eHealth (3): a systematic review of published definitions. Journal of Medical Internet Research. 2005;7(1):e1-e1. DOI: 10.2196/jmir.7.1.e1.
- [5] World Health Organization (WHO). eHealth at WHO. [cited 07.08.2020]. Available from: <https://www.who.int/ehealth/en/>.
- [6] Lunde P., Nilsson B. B., Bergland A., Kværner K. J. and Bye A. The Effectiveness of Smartphone Apps for Lifestyle Improvement in Noncommunicable Diseases: Systematic Review and Meta-Analyses. Journal of Medical Internet Research. 2018;20(5):e162-e162. DOI: 10.2196/jmir.9751.
- [7] Wang K., Varma D. S. and Prosperi M. A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. Journal of Psychiatric Research. 2018;107:73-78. DOI: <https://doi.org/10.1016/j.jpsychires.2018.10.006>.
- [8] Lampariello R. and Labrique A. B. Digital health: Is the glass half-full or half-empty? Health Policy and Technology. 2020;9(3):266-267. DOI: <https://doi.org/10.1016/j.hlpt.2020.07.001>.
- [9] Albrecht U. V. H. and Medizinische Hochschule Hannover (MHH). Chancen und Risiken von Gesundheits-Apps (CHARISMHA); engl. Chances and Risks of Mobile Health Apps (CHARISMHA). 2016. [cited 04.10.2020]. Available from: <http://www.digibib.tu-bs.de/?docid=00060000>.
- [10] Neumann K., Larisch K., Dietzel J., Kurepkat M., Weißer W. and Wenzlau V. Digital healthcare products – Leveraging opportunities – developing safe routes to market. 2016 [cited 06.08.2020]. Available from: https://www.iges.com/e6/e1621/e10211/e13470/e15278/e15279/e15281/attr_objs15282/IGES_Publication_Digital_healthcare_products_WEB_eng.pdf.
- [11] Bundesministerium für Gesundheit. Digitale Versorgung Gesetz – DVG. 2020 [cited 07.08.2020]. Available from: <https://www.bundesgesundheitsministerium.de/digitale-versorgung-gesetz.html>.
- [12] Moshi M. R., Tooher R. and Merlin T. Suitability of Current Evaluation Frameworks for Use in the Health Technology Assessment of Mobile Medical Applications: A Systematic Review. International Journal of Technology Assessment in Health Care. 2018;34(5):464-475.
- [13] Vis C., Buhrmann L., Riper H. and Ossebaard H. C. Health technology assessment frameworks for eHealth: A systematic review. International Journal of Technology Assessment in Health Care. 2020;1-13.
- [14] Expert Panel on effective ways of investing in health (EXPH). Assessing the Impact of Digital Transformation of Health Services. 2019. [cited 10.09.2020]. Available from: https://ec.europa.eu/health/sites/health/files/expert_panel/docs/022_digitaltransformation_en.pdf.
- [15] World Health Organization (WHO). Classification of digital health interventions v1.0, A shared language to describe the uses of digital technology for health. 2018. Available from: <https://www.who.int/reproductivehealth/publications/mhealth/classification-digital-health-interventions/en/>.
- [16] World Health Organization (WHO). WHO Guideline: recommendations on digital interventions for health system strengthening 2019.
- [17] 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.

- [18] European Commission (EC). Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. Official Journal of the European Union. 2017;L 117/1.
- [19] European Commission (EC). Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU. Official Journal of the European Union. 2017;L 117/176.
- [20] 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 10.09.2020]. Available from: <https://ec.europa.eu/docsroom/documents/37581>.
- [21] Medical Device Coordination Group (MDCG). Guidance on Clinical Evaluation (MDR)/Performance Evaluation (IVDR) of Medical Device Software. 2020. [cited 10.09.2020]. Available from: <https://ec.europa.eu/docsroom/documents/40323>.
- [22] International Medical Device Regulators Forum (IMDRF). Software as a Medical Device (SaMD): Clinical Evaluation. 2017. [cited 10.09.2020]. Available from: <http://www.imdrf.org/consultations/cons-samd-ce.asp>.
- [23] Ouzzani M., Hammady H., Fedorowicz Z. and Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews*. 2016;5(1):210. DOI: 10.1186/s13643-016-0384-4.
- [24] Med Tech Reimbursement Consulting (MTRC). Reimbursement Landscape for Health Apps in Europe. 2020.
- [25] Hoffmann T. C., Glasziou P. P., Boutron I., Milne R., Perera R., Moher D., et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ: British Medical Journal*. 2014;348:g1687. DOI: 10.1136/bmj.g1687.
- [26] Moshi M. R., Parsons J., Tooher R. and Merlin T. Evaluation of Mobile Health Applications: Is Regulatory Policy Up to the Challenge? *International Journal of Technology Assessment in Health Care*. 2019;35(4):351-360.
- [27] Moshi M. R., Tooher R. and Merlin T. Development of a health technology assessment module for evaluating mobile medical applications. *International Journal of Technology Assessment in Health Care*. 2020:1-10.
- [28] National Institute for Health and Care Excellence (NICE). Evidence Standards Framework for Digital Health Technologies. 2019 [cited 07.08.2020]. Available from: <https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/evidence-standards-framework/digital-evidence-standards-framework.pdf>.
- [29] Sauerland S. Digitale Gesundheitsinterventionen – Wann braucht es welche Evidenz?: 2019 [cited 07.08.2020]. Available from: https://www.ebm-netzwerk.de/de/medien/pdf/ebm-04_19_kvh_journal_digitale-inventionen.pdf.
- [30] 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(4). DOI: 10.23996/fjhw.82538.
- [31] Collins L. M., Murphy S. A. and Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. *American journal of preventive medicine*. 2007;32(5 Suppl):S112-S118. DOI: 10.1016/j.amepre.2007.01.022.
- [32] Kowatsch T., Otto L., Harperink S., Cotti A. and Schlieter H. A design and evaluation framework for digital health interventions. *it – Information Technology*. 2019;61(5-6):253-263. DOI: 10.1515/itit-2019-0019.
- [33] Mathews S. C., McShea M. J., Hanley C. L., Ravitz A., Labrique A. B. and Cohen A. B. Digital health: a path to validation. *Npj Digital Medicine*. 2019;2:38.
- [34] Wilson K., Bell C., Wilson L. and Witteman H. Agile research to complement agile development: a proposal for an mHealth research lifecycle. *Npj Digital Medicine*. 2018;1:46.

- [35] Shaw N. T. 'CHEATS': a generic information communication technology (ICT) evaluation framework. *Comput Biol Med.* 2002;32(3):209-220. Epub 2002/03/30. DOI: 10.1016/s0010-4825(02)00016-1.
- [36] Kidholm K., Ekeland A. G., Jensen L. K., Rasmussen J., Pedersen C. D., Bowes A., et al. A MODEL FOR ASSESSMENT OF TELEMEDICINE APPLICATIONS: MAST. *International Journal of Technology Assessment in Health Care.* 2012;28(1):44-51. Epub 2012/01/23. DOI: 10.1017/S0266462311000638.
- [37] Currie W. TEMPEST: An integrative model for health technology assessment. *Health Policy and Technology.* 2012;1:35-49. DOI: 10.1016/j.hlpt.2012.01.004.
- [38] Stoyanov S. R., Hides L., Kavanagh D. J., Zelenko O., Tjondronegoro D. and Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR MHealth and UHealth.* 2015;3(1):e27-e27. DOI: 10.2196/mhealth.3422.
- [39] Mummah S. A., Robinson T. N., King A. C., Gardner C. D. and Sutton S. IDEAS (Integrate, Design, Assess, and Share): A Framework and Toolkit of Strategies for the Development of More Effective Digital Interventions to Change Health Behavior. *Journal of Medical Internet Research.* 2016;18(12):e317-e317. DOI: 10.2196/jmir.5927.
- [40] Eivazzadeh S., Anderberg P., Larsson T. C., Fricker S. A. and Berglund J. Evaluating Health Information Systems Using Ontologies. *JMIR medical informatics.* 2016;4(2):e20-e20. DOI: 10.2196/medinform.5185.
- [41] Leigh S., Ouyang J. and Mimmagh C. Effective? Engaging? Secure? Applying the ORCHA-24 framework to evaluate apps for chronic insomnia disorder. *Evidence Based Mental Health.* 2017;20(4):e20. DOI: 10.1136/eb-2017-102751.
- [42] Zhou L., Bao J., Setiawan I. M. A., Saptono A. and Parmanto B. The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study. *JMIR MHealth and UHealth.* 2019;7(4):e11500.
- [43] American Psychiatric Association. App Evaluation Model. 2020 [cited 06.08.2020]. Available from: <https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/the-app-evaluation-model>.
- [44] Jeminiwa R. N., Hohmann N. S. and Fox B. I. Developing a Theoretical Framework for Evaluating the Quality of mHealth Apps for Adolescent Users: A Systematic Review. *The Journal of Pediatric Pharmacology & Therapeutics.* 2019;24(4):254-269.
- [45] Henson P., David G., Albright K. and Torous J. Deriving a practical framework for the evaluation of health apps. *The Lancet Digital Health.* 2019;1(2):e52-e54. DOI: 10.1016/s2589-7500(19)30013-5.
- [46] Enam A., Torres-Bonilla J. and Eriksson H. Evidence-Based Evaluation of eHealth Interventions: Systematic Literature Review. *Journal of Medical Internet Research.* 2018;20(11):e10971.
- [47] Kolasa K. and Kozinski G. How to Value Digital Health Interventions? A Systematic Literature Review. *International Journal of Environmental Research & Public Health [Electronic Resource].* 2020;17(6):23.
- [48] Holl F. and Swoboda W. Methods to Measure the Impact of mHealth Applications: Preliminary Results of a Scoping Review. *Studies in Health Technology & Informatics.* 2018;251:285-288.
- [49] Maier T., Kulichova D., Schotten K., Astrid R., Ruzicka T., Berking C., et al. Accuracy of a smartphone application using fractal image analysis of pigmented moles compared to clinical diagnosis and histological result. *Journal of the European Academy of Dermatology and Venereology.* 2015;29(4):663-667. DOI: 10.1111/jdv.12648.
- [50] Thissen M., Udrea A., Hacking M., von Braunmuehl T. and Ruzicka T. mHealth App for Risk Assessment of Pigmented and Nonpigmented Skin Lesions—A Study on Sensitivity and Specificity in Detecting Malignancy. *Telemedicine and e-Health.* 2017;23(12):948-954. DOI: 10.1089/tmj.2016.0259.
- [51] Udrea A., Mitra G. D., Costea D., Noels E. C., Wakkee M., Siegel D. M., et al. Accuracy of a smartphone application for triage of skin lesions based on machine learning algorithms. *Journal of the European Academy of Dermatology and Venereology.* 2020;34(3):648-655. DOI: 10.1111/jdv.15935.
- [52] de Carvalho T. M., Noels E., Wakkee M., Udrea A. and Nijsten T. Development of Smartphone Apps for Skin Cancer Risk Assessment: Progress and Promise. *JMIR Dermatol.* 2019;2(1):e13376. DOI: 10.2196/13376.

- [53] 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.
- [54] Mindable Health UG. Mindable Wissenschaft. [cited 03.09.2020]. Available from: <https://www.mindable.health/wissenschaft/>.
- [55] Berglund Scherwitzl E., Gemzell Danielsson K., Sellberg J. A. and Scherwitzl R. Fertility awareness-based mobile application for contraception. *The European Journal of Contraception & Reproductive Health Care*. 2016;21(3):234-241. DOI: 10.3109/13625187.2016.1154143.
- [56] Berglund Scherwitzl E., Lundberg O., Kopp Kallner H., Gemzell Danielsson K., Trussell J. and Scherwitzl R. Perfect-use and typical-use Pearl Index of a contraceptive mobile app. *Contraception*. 2017;96(6):420-425. DOI: <https://doi.org/10.1016/j.contraception.2017.08.014>.
- [57] Bull J., Rowland S., Lundberg O., Berglund-Scherwitzl E., Gemzell-Danielsson K., Trussell J., et al. Typical use effectiveness of Natural Cycles: postmarket surveillance study investigating the impact of previous contraceptive choice on the risk of unintended pregnancy. *BMJ Open*. 2019;9(3):e026474. DOI: 10.1136/bmjopen-2018-026474.
- [58] Berglund Scherwitzl E., Lundberg O., Kopp Kallner H., Rowland S. P., Holte J., Trussell J., et al. Short- and long-term effect of contraceptive methods on fecundity. *The European Journal of Contraception & Reproductive Health Care*. 2019;24(4):260-265. DOI: 10.1080/13625187.2019.1621999.
- [59] Bull J. R., Rowland S. P., Scherwitzl E. B., Scherwitzl R., Danielsson K. G. and Harper J. Real-world menstrual cycle characteristics of more than 600,000 menstrual cycles. *Npj Digital Medicine*. 2019;2:83-83. DOI: 10.1038/s41746-019-0152-7.
- [60] Kleinschmidt T. K., Bull J. R., Lavorini V., Rowland S. P., Pearson J. T., Scherwitzl E. B., et al. Advantages of determining the fertile window with the individualised Natural Cycles algorithm over calendar-based methods. *The European Journal of Contraception & Reproductive Health Care*. 2019;24(6):457-463. DOI: 10.1080/13625187.2019.1682544.
- [61] Koenig N., Seeck A., Eckstein J., Mainka A., Huebner T., Voss A., et al. Validation of a New Heart Rate Measurement Algorithm for Fingertip Recording of Video Signals with Smartphones. *Telemedicine and e-Health*. 2016;22(8):631-636. DOI: 10.1089/tmj.2015.0212.
- [62] Krivoshei L., Weber S., Burkard T., Maseli A., Brasier N., Kühne M., et al. Smart detection of atrial fibrillation†. *Europace: European pacing, arrhythmias, and cardiac electrophysiology: journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology*. 2017;19(5):753-757. DOI: 10.1093/europace/euw125.
- [63] Weber S., Eckstein J., Koenig N., Seeck A. and Birkemeyer R. P1617Detection of premature beats using smartphone-based pulse wave recordings. *EP Europace*. 2017;19(suppl_3):iii347-iii347. DOI: 10.1093/ehjci/eux158.243.
- [64] Karim N., Aral M., Eduawor S. and Finlay M. AF detection using smartphone apps. *Europace*. 2017;19:iii234-iii234. DOI: 10.1093/ehjci/eux151.246.
- [65] Brasier N., Raichle C. J., Dörr M., Becke A., Nohturfft V., Weber S., et al. Detection of atrial fibrillation with a smartphone camera: first prospective, international, two-centre, clinical validation study (DETECT AF PRO). *EP Europace*. 2019;21(1):41-47. DOI: 10.1093/europace/euy176.
- [66] Dörr M., Nohturfft V., Brasier N., Bosshard E., Djurdjevic A., Gross S., et al. The WATCH AF Trial: SmartWATCHes for Detection of Atrial Fibrillation. *JACC: Clinical Electrophysiology*. 2019;5(2):199-208. DOI: <https://doi.org/10.1016/j.jacep.2018.10.006>.
- [67] Birkemeyer R., Müller A., Wahler S. and von der Schulenburg J.-M. A cost-effectiveness analysis model of Preventicus atrial fibrillation screening from the point of view of statutory health insurance in Germany. *Health Economics Review*. 2020;10(1):16. DOI: 10.1186/s13561-020-00274-z.
- [68] Huber S., Priebe J. A., Baumann K.-M., Plidschun A., Schiessl C. and Tölle T. R. Treatment of Low Back Pain with a Digital Multidisciplinary Pain Treatment App: Short-Term Results. *JMIR Rehabil Assist Technol*. 2017;4(2):e11. DOI: 10.2196/rehab.9032.

- [69] Clement I., Lorenz A., Ulm B., Plidschun A. and Huber S. Implementing Systematically Collected User Feedback to Increase User Retention in a Mobile App for Self-Management of Low Back Pain: Retrospective Cohort Study. *JMIR Mhealth Uhealth*. 2018;6(6):e10422. DOI: 10.2196/10422.
- [70] Toelle T. R., Utpadel-Fischler D. A., Haas K.-K. and Priebe J. A. App-based multidisciplinary back pain treatment versus combined physiotherapy plus online education: a randomized controlled trial. *Npj Digital Medicine*. 2019;2(1):34. DOI: 10.1038/s41746-019-0109-x.
- [71] Scherr S. and Goering M. Is a Self-Monitoring App for Depression a Good Place for Additional Mental Health Information? Ecological Momentary Assessment of Mental Help Information Seeking among Smartphone Users. *Health Communication*. 2019;35:1-9. DOI: 10.1080/10410236.2019.1606135.
- [72] Martinengo L., Van Galen L., Lum E., Kowalski M., Subramaniam M. and Car J. Suicide prevention and depression apps' suicide risk assessment and management: a systematic assessment of adherence to clinical guidelines. *BMC medicine*. 2019;17(1):231-231. DOI: 10.1186/s12916-019-1461-z.
- [73] Okamoto H., Stracke H., Stoll W. and Pantev C. Listening to tailor-made notched music reduces tinnitus loudness and tinnitus-related auditory cortex activity. *Proceedings of the National Academy of Sciences*. 2010;107(3):1207. DOI: 10.1073/pnas.0911268107.
- [74] Stein A., Wunderlich R., Lau P., Engell A., Wollbrink A., Shaykevich A., et al. Clinical trial on tonal tinnitus with tailor-made notched music training. *BMC Neurology*. 2016;16(1):38. DOI: 10.1186/s12883-016-0558-7.
- [75] Löhler J. and Dobel C. Tinnitracks – Erste Ergebnisse einer Anwenderbefragung. *Laryngorhinootologie*. 2019;98(11):789-796. Epub 26.09.2019. DOI: 10.1055/a-1007-1858.
- [76] Miller A. S., Cafazzo J. A. and Seto E. A game plan: Gamification design principles in mHealth applications for chronic disease management. *Health Informatics Journal*. 2014;22(2):184-193. DOI: 10.1177/1460458214537511.
- [77] Peng W., Yuan S. and Holtz B. E. Exploring the Challenges and Opportunities of Health Mobile Apps for Individuals with Type 2 Diabetes Living in Rural Communities. *Telemedicine and e-Health*. 2016;22(9):733-738. DOI: 10.1089/tmj.2015.0180.
- [78] Martinez M., Park S. B., Maison I., Mody V., Soh L. S. and Parihar H. S. iOS Appstore-Based Phone Apps for Diabetes Management: Potential for Use in Medication Adherence. *JMIR Diabetes*. 2017;2(2):e12. DOI: 10.2196/diabetes.6468.
- [79] Kebede M. M. and Pischke C. R. Popular Diabetes Apps and the Impact of Diabetes App Use on Self-Care Behaviour: A Survey Among the Digital Community of Persons With Diabetes on Social Media. *Frontiers in Endocrinology*. 2019;10:135.
- [80] Fu H. N. C., Adam T. J., Konstan J. A., Wolfson J. A., Clancy T. R. and Wyman J. F. Influence of Patient Characteristics and Psychological Needs on Diabetes Mobile App Usability in Adults With Type 1 or Type 2 Diabetes: Crossover Randomized Trial. *JMIR Diabetes*. 2019;4(2):e11462. DOI: 10.2196/11462.
- [81] Debong F., Mayer H. and Kober J. Real-World Assessments of mySugr Mobile Health App. *Diabetes Technology & Therapeutics*. 2019;21(S2):S2-35-S32-40. DOI: 10.1089/dia.2019.0019.
- [82] Quevedo Rodríguez A. and Wägner A. M. Mobile phone applications for diabetes management: A systematic review. *Endocrinología, Diabetes y Nutrición*. 2019;66(5):330-337. DOI: <https://doi.org/10.1016/j.endinu.2018.11.005>.
- [83] Mayer H., Bankosegger R. P. and Kober J. 953-P: Sustainable Improvement in Quality of Blood Glucose Control in Users of mySugr's Integrated Diabetes Management Solution. *Diabetes*. 2019;68(Supplement 1):953-P. DOI: 10.2337/db19-953-P.
- [84] Fu H. N. C., Rizvi R. F., Wyman J. F. and Adam T. J. Usability Evaluation of Four Top-Rated Commercially Available Diabetes Apps for Adults With Type 2 Diabetes. *CIN: Computers, Informatics, Nursing*. 2020;38(6).
- [85] Freeman K., Dinnes J., Chuchu N., Takwoingi Y., Bayliss S. E., Matin R. N., et al. Algorithm based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies. *BMJ*. 2020;368:m127. DOI: 10.1136/bmj.m127.

- [86] Bashi N., Fatehi F., Mosadeghi-Nik M., Askari M. S. and Karunanithi M. Digital health interventions for chronic diseases: a scoping review of evaluation frameworks. *BMJ Health & Care Informatics*. 2020;27(1).
- [87] Moshi M., Tooher R. and Merlin T. mHealth app evaluation framework for reimbursement decision-making. *International Journal of Technology Assessment in Health Care*. 2019;35:33.
- [88] Gordon W. J., Landman A., Zhang H. and Bates D. W. Beyond validation: getting health apps into clinical practice. *NPJ Digit Med*. 2020;3:14. Epub 2020/02/13. DOI: 10.1038/s41746-019-0212-z.
- [89] MHealthBelgium. Validation Pyramid. 2020. [cited 17.09.2020]. Available from: <https://mhealthbelgium.be/validation-pyramid>.
- [90] Med Tech Reimbursement Consulting (MTRC). The first mobile app reimbursed in Belgium. 2020. [cited 17.09.2020]. Available from: https://mtrconsult.com/news/first-mobile-app-reimbursed-belgium?utm_source=Newsletter&utm_campaign=2964e8ddf2-EMAIL_CAMPAIGN_2020_08_03_08_30&utm_medium=email&utm_term=0_922cb68711-2964e8ddf2-134005135.
- [91] Haute Autorité de Santé (HAS). Medical device evaluation by the CNEDiMTS (Medical Device and Health Technology Evaluation Committee): Guide to the specific features of clinical evaluation of a connected medical device (CMD) in view of its application for reimbursement. 2019.
- [92] Med Tech Reimbursement Consulting (MTRC). First health app in the French LPPR list for add-on reimbursement. 2020. [cited 17.09.2020]. Available from: <https://mtrconsult.com/news/first-health-app-french-lppr-list-add-reimbursement>.
- [93] Martelli N. and van den Brink H. Special funding schemes for innovative medical devices in French hospitals: The pros and cons of two different approaches. *Health Policy*. 2014;117(1):1-5 (<https://doi.org/10.1016/j.healthpol.2014.1004.1007>).
- [94] Haute Autorité de Santé (HAS). Assessment and Improvement of Practice: Good Practice Guidelines on Health Apps and Smart Devices (Mobile Health or mHealth). 2016. [cited 17.09.2020]. Available from: https://www.has-sante.fr/upload/docs/application/pdf/2017-03/dir1/good_practice_guidelines_on_health_apps_and_smart_devices_mobile_health_or_mhealth.pdf.
- [95] Med Tech Reimbursement Consulting (MTRC). Reimbursement of telemedicine and digital care in the Netherlands. 2020. [cited 17.09.2020]. Available from: <https://mtrconsult.com/news/reimbursement-telemedicine-and-digital-care-netherlands>.
- [96] National eHealth Living Lab (NeLL). NeLL projects criteria. 2020. [cited 17.09.2020]. Available from: <https://nell.eu/english#criteria>.
- [97] MedTech Reimbursement Consulting. Reimbursement Landscape for Health Apps in Europe. 2020. [cited 17.09.2020]. Available from: <https://www.a-marketaccess.com/healthapps>.
- [98] Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM). Das Fast Track Verfahren für digitale Gesundheitsanwendungen (DiGA). 2020.
- [99] Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM). BfArM nimmt erste „Apps auf Rezept“ ins Verzeichnis digitaler Gesundheitsanwendungen (DiGA) auf. 4/20 ed2020.
- [100] Institut für Gesundheitsförderung und Prävention (IfGP). Gesundheits-Apps, Grundlagenpapier unter besonderer Berücksichtigung des Aspekts Gesundheitskompetenz. 2018. [cited 17.09.2020]. Available from: <https://oepgk.at/wp-content/uploads/2018/10/grundlagenpapier-gesundheits-apps.pdf>.
- [101] Knöppler K. and al. e. Digital-Health-Anwendungen für die Versorgung nutzbar machen (Teil 1 bis Teil 6). 2016-2019.
- [102] Böhler C. Methodische Empfehlungen für die ökonomische Evaluation von eHealth-Applikationen in Österreich 2018.. 2018. [cited 17.09.2020]. Available from: <https://www.sozialversicherung.at/cdscontent/load?contentid=10008.712915&version=1533548852>.
- [103] Wackerle H. Der digitale Gesundheitspfad. In: ITSv, editor.2020.
- [104] Angelescu K. and Sauerland S. Mobile Gesundheitsanwendungen: Welche Evidenz ist nötig? *Dtsch Arztebl*. 2019;116(21):A-1057/B-1870/C-1858.

- [105] Bradway M., Carrion C., Vallespin B., Saadatfard O., Puigdomènech E., Espallargues M., et al. mHealth Assessment: Conceptualization of a Global Framework. *JMIR Mhealth Uhealth*. 2017;5(5):e60. DOI: 10.2196/mhealth.7291.
- [106] European Commission (EC). Green paper on mobile health („mHealth“). 2014 [cited 04.10.2020]. Available from: <https://ec.europa.eu/digital-single-market/en/news/green-paper-mobile-health-mhealth>.
- [107] TK legt Kooperation mit Ada Health auf Eis. *Ärzte Zeitung online: Springer Medizin Verlag GmbH*; 2019.
- [108] Babylon Health data breach: GP app users able to see other people’s consultations. *The Guardian* 2020.
- [109] Evans J., Bhatt S., Sharma R., Baulch E., Watkins J., Tariq A., et al. The Path to Scale: Navigating Design, Policy, and Infrastructure. Evans, Jay. *Medic Mobile*, Kathmandu, Nepal Evans, Jay. *Global Health Academy*, University of Edinburgh, Edinburgh, UK Bhatt, Shreya. *Medic Mobile*, Mumbai, India Sharma, Ranju. *Medic Mobile*, Kathmandu, Nepal Baulch, Emma. *Digital Media Research Centre*, Queensland University of Technology, Brisbane, Queensland, Australia Watkins, Jerry. *School of Communication and Design*, RMIT University Vietnam, Ho Chi Minh City, Vietnam Tariq, Amina. *School of Public Health and Social Work*, Queensland University of Technology, Brisbane, Queensland, Australia Dordrecht: Springer; 2018 2018. 31-48 p.
- [110] Bradway M., Gabarron E., Johansen M., Zanaboni P., Jardim P., Joakimsen R., et al. Methods and Measures Used to Evaluate Patient-Operated Mobile Health Interventions: Scoping Literature Review. *JMIR MHealth and UHealth*. 2020;8(4):e16814.
- [111] Liu X., Faes L., Kale A. U., Wagner S. K., Fu D. J., Bruynseels A., et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *The Lancet Digital Health*. 2019;1(6):e271-e297. DOI: [https://doi.org/10.1016/S2589-7500\(19\)30123-2](https://doi.org/10.1016/S2589-7500(19)30123-2).
- [112] Deutscher Ethikrat. Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung. 2017 [cited 04.10.2020]. [cited 04.10.2020]. Available from: <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-big-data-und-gesundheit.pdf>.
- [113] Liu X., Cruz Rivera S., Moher D., Calvert M. J., Denniston A. K., Chan A.-W., et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nature Medicine*. 2020;26(9):1364-1374. DOI: 10.1038/s41591-020-1034-x.
- [114] Haasteren A., Gille F., Fadda M. and Vayena E. Development of the mHealth App Trustworthiness checklist. *DIGITAL HEALTH*. 2019;5:205520761988646. DOI: 10.1177/2055207619886463.
- [115] Kramer U. HealthOn Checkliste für Gesundheits-Apps. 2018 [cited 04.10.2020]. [cited 04.10.2020]. Available from: <https://www.healthon.de/checkliste>.
- [116] Tønning M., Kessing L., Bardram J. and Faurholt-Jepsen M. Methodological Challenges in Randomized Controlled Trials on Smartphone-Based Treatment in Psychiatry: Systematic Review. *J Med Internet Res*. 2019;21(10):e15362. Epub 2019/10/31. DOI: 10.2196/15362.
- [117] Pham Q., Wiljer D. and Cafazzo J. Beyond the Randomized Controlled Trial: A Review of Alternatives in mHealth Clinical Trial Methods. *JMIR Mhealth Uhealth*. 2016;4(3):e107. Epub 2016/09/11. DOI: 10.2196/mhealth.5720.
- [118] Klasnja P., Hekler E., Shiffman S., Boruvka A., Almirall D., Tewari A., et al. Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. *Health Psychol*. 2015;34s(0):1220-1228. Epub 2015/12/15. DOI: 10.1037/hea0000305.
- [119] Walton A., Nahum-Shani I., Crosby L., Klasnja P. and Murphy S. Optimizing Digital Integrated Care via Micro-Randomized Trials. *Clinical Pharmacology & Therapeutics*. 2018;104(1):53-58. DOI: 10.1002/cpt.1079.
- [120] Liao P., Klasnja P., Tewari A. and Murphy S. Micro-Randomized Trials in mHealth. 2015 Report No. arXiv:1504.00238v1.
- [121] Liao P., Klasnja P., Tewari A. and Murphy S. Sample size calculations for micro-randomized trials in mHealth. *Stat Med*. 2016;35(12):1944-1971. Epub 2015/12/29. DOI: 10.1002/sim.6847.

- [122] Lei H., Nahum-Shani I., Lynch K., Oslin D. and Murphy S. A „SMART“ design for building individualized treatment sequences. *Annu Rev Clin Psychol.* 2012;8:21-48. Epub 2012/01/10. DOI: 10.1146/annurev-clinpsy-032511-143152.
- [123] Shcherbatykh I., Holbrook A., Thabane L. and Dolovich L. Methodologic issues in health informatics trials: the complexities of complex interventions. *J Am Med Inform Assoc.* 2008;15(5):575-580. Epub 2008/06/27. DOI: 10.1197/jamia.M2518.

7 Anhang

7.1 mHealth Assessment Frameworks

7.1.1 HTA Module for Mobile Medical Applications (MMAs)

Table 7-1: A module which can be used in addition to existing core HTA domains to adapt HTA and/or reimbursement evaluation frameworks to assess MMAs [27]

HTA domain	HTA domain status	MMA technology-specific considerations	
		Challenge posed by MMA	MMA specific modifications and adaptation(s)
Description and technical characteristics	Mandatory	Operating system(s) for MMA	Operating systems of the MMA (i.e. Android, iOS, etc.)
		Operating platform(s) for MMA	Operating platforms of the MMA (i.e. smartphone, tablet, smartwatch)
Current use of the technology		Rationale for use	- The intended purpose of the MMA (i.e. diagnose, treat, inform clinical management, clinical management) - The healthcare condition or situation that the MMA addresses - MMA input (i.e. image, physiological status, symptoms, etc.) - MMA algorithm (i.e. equations, analysis engine model logic, algorithm, etc.) - MMA output (i.e. inform, treat, diagnose)
		Potential software changes (i.e. updates)	- Post-market software changes, that do not require the re-evaluation of an MMA and are corrective, preventive, adaptive, and/or perfective (see <i>Effectiveness</i> for more information) - Post-market software changes that do require a re-evaluation of the effectiveness and safety of an MMA and which enable or disable new MMA functions (see <i>Effectiveness</i> for more information)
Effectiveness		Accuracy (i.e. constant error)	- Closeness of the output to the true value to the MMA's output - Accuracy measures the effect of software errors on the MMA output - Example: In psychometrics, accuracy is the degree to which test scores are supported by evidence and theory.
		Configuration	- The MMA's ability to withstand user configuration in an unintended way (i.e. results of <i>fuzzing</i> or <i>fuzz testing</i>) - Limitations of the MMA (i.e. assumptions, data quality, algorithms)
		Communication and display	- The design of the MMA user interface (i.e. level of complexity, type of platform, how information is displayed, etc.) - The appropriateness of the MMA interface as a means of information display (i.e. language translation, units displayed, clarity, etc.) - The MMA's ability to communicate the relevant information (i.e. data quality, network availability, correct installation, etc.)
		Cybersecurity and connectivity	- Formalised and safe methods have been implemented to convert, transmit and/or store MMA data (i.e. results of <i>fuzzing</i> or <i>fuzz testing</i>) - Users can safely implement information security updates - System supports ensure protection of MMA system information - MMA software adheres to robust programming principles (i.e. paranoia, stupidity, dangerous implements, cannot happen) - Balances the availability of timely information and against privacy and security (i.e. results of <i>fuzzing</i> or <i>fuzz testing</i>)

HTA domain	HTA domain status	MMA technology-specific considerations	
		Challenge posed by MMA	MMA specific modifications and adaptation(s)
		Cybersecurity and connectivity (continuation)	- How MMA integrates with other software (i.e. results of <i>fuzzing</i> or <i>fuzz testing</i>) - The need for MMA security software to be updated so that it can be used alongside other systems, applications or in operating environments (i.e. results of <i>fuzzing</i> or <i>fuzz testing</i>)
		Potential software changes (i.e. updates)	- Adaptive software changes (i.e. maintains software with dynamic environment) - Perfective software changes (i.e. recoding to improve performance) - Corrective software changes (i.e. corrects problems) - Preventive software changes (i.e. corrects latent faults before they cause operational problems)
		Precision (i.e. variable error)	- Under unchanged conditions, the degree to which the recurrent measurements input into the MMA generates the same output (e.g. reproducibility, consistency). - Example: in psychology precision is degree in which two or more of measurement of the same tool consistently arrive at the same result.
		Analytical validity	- MMA's ability to reliably and accurately produce the intentional output from the input data - The algorithm used by the MMA is a recognised standard (the current standard of care or described in the literature (i.e. insulin dosing)) - MMA accuracy is relative to reference standard (i.e. <i>International normalisation ratio (INR)</i>) - MMA comparable to another software or device that has an association between the software output and a health outcome
Safety		The risk of misinformation	How the MMA output (i.e. information) affects clinical decision making regarding management of a patient's condition
Cost-Effectiveness (in fee-for-service model)		Technological evolution	- Considerations of applicability of the system, platform, licensing, attachable hardware, and versions of the MMA to those that would be used in the health system - Unit costs including MMA costs and in-app purchases
Organisational aspects		Digital health literacy	- The training/education (i.e. digital literacy) which may be needed for the user(s) (i.e. medical practitioners, patients, caregivers) to effectively utilise the MMA. Examples include: - Continual professional development (CPD) courses for medical practitioner(s) to effectively learn how to utilise and recommend MMA's in clinical practice - Education that patient(s) have to undergo to effectively learn how to utilise MMA's
		Responsibility	Accreditation that may be needed for professionals (i.e. medical practitioners, allied health workers, technicians) to prescribe and/or use the MMA
		Connectivity	The MMA interaction with current health informatics systems (i.e. hospitals and surgeries). Examples of health informatics systems include, but are not limited to, <i>PROCURA</i> and <i>Enterprise patient administration system (EPAS)</i>
		Technological evolution	- How adopting the MMA could alter the current utilisation of services (i.e. workload, workforce, compliance, etc.) - How adopting the MMA will change treatment location (i.e. home-based, rural, remote, hospital, clinic, etc.)
Ethical aspects	Patient privacy and confidentiality	- The presence of a privacy policy - The contents of a privacy policy	
	Equity concerns	Considerations include user disability (how could users' with blindness use the MMA), language (users' who have English as a second language), age, literacy, socio-economic status, etc.	
	Access concerns	Considerations include the cost of platform, in-app purchases, cost of MMA, geographical location, internet availability etc.	
	Technological evolution	Any possible conflicts of interest (i.e. developer or owner affiliation, sources funding, third-party sponsorship, etc.)	

HTA domain	HTA domain status	MMA technology-specific considerations	
		Challenge posed by MMA	MMA specific modifications and adaptation(s)
Legal aspects		Responsibility	■ Litigation risks to the relevant person(s) associated with the use or recommendation of the MMA for healthcare practitioners (i.e. GPs, allied health workers, etc.) ■ How insurance(s) (i.e. professional indemnity, life, health, income) for all stakeholders (i.e. patients, medical professionals, developers) could be affected through use or recommendation of the MMA ■ How possible professional registrations could be affected through the use or recommendation of the MMA (e.g. for medical practitioners with AHPRA) ■ Clarify which party owns the data related to the MMA (i.e. patient, third party, medical practitioners) ■ Clarity around which party (i.e. manufacturer, medical practitioner who prescribed it) is responsible for the medical advice provided by the MMA ■ Clarity around which party (i.e. manufacturer, medical practitioner, app developer) is responsible for monitoring and reviewing the patient data entered into the MMA
Post-market monitoring		Reappraisal	■ Post-market data that requires a full review of the effectiveness and safety of an MMA. These are performance data that alter the effective measures of the MMA (i.e. inferior or superior to the original measures stated in the original HTA) and/or which change the harms posed by the MMA (i.e. inferior or superior to the original measures stated in the original HTA) ■ How the manufacturer plans to monitor the MMA's performance data (i.e. data includes user feedback, complaints, and adverse events, real-world data, etc.) ■ How the data collection implemented has the least user burdensome approach to collect the MMA's performance data ■ How the post-market data could be used to enable or disable new MMA functionalities (i.e. addition or removal of functionalities stated in the original submission, etc.) ■ How post-market data could affect the MMA's cost-effectiveness, safety, effectiveness ■ How post-market data could affect the ethical, legal, and/or organisational concerns associated with the MMA
Social aspects	Optional	None	■ How the use of the MMA may affect the patients' caregiver(s), including relationships with medical professionals ■ How the use of the MMA may affect the users' relationships (i.e. family dynamics, friends, and other relevant social relations) ■ How the MMA may benefit patient autonomy

APHRA – Australian Practitioner Regulation Agency; GP – General practitioner; HTA – Health technology assessment; MMA – Mobile medical applications.

7.1.2 Evidence Standards Framework for Digital Health Technologies

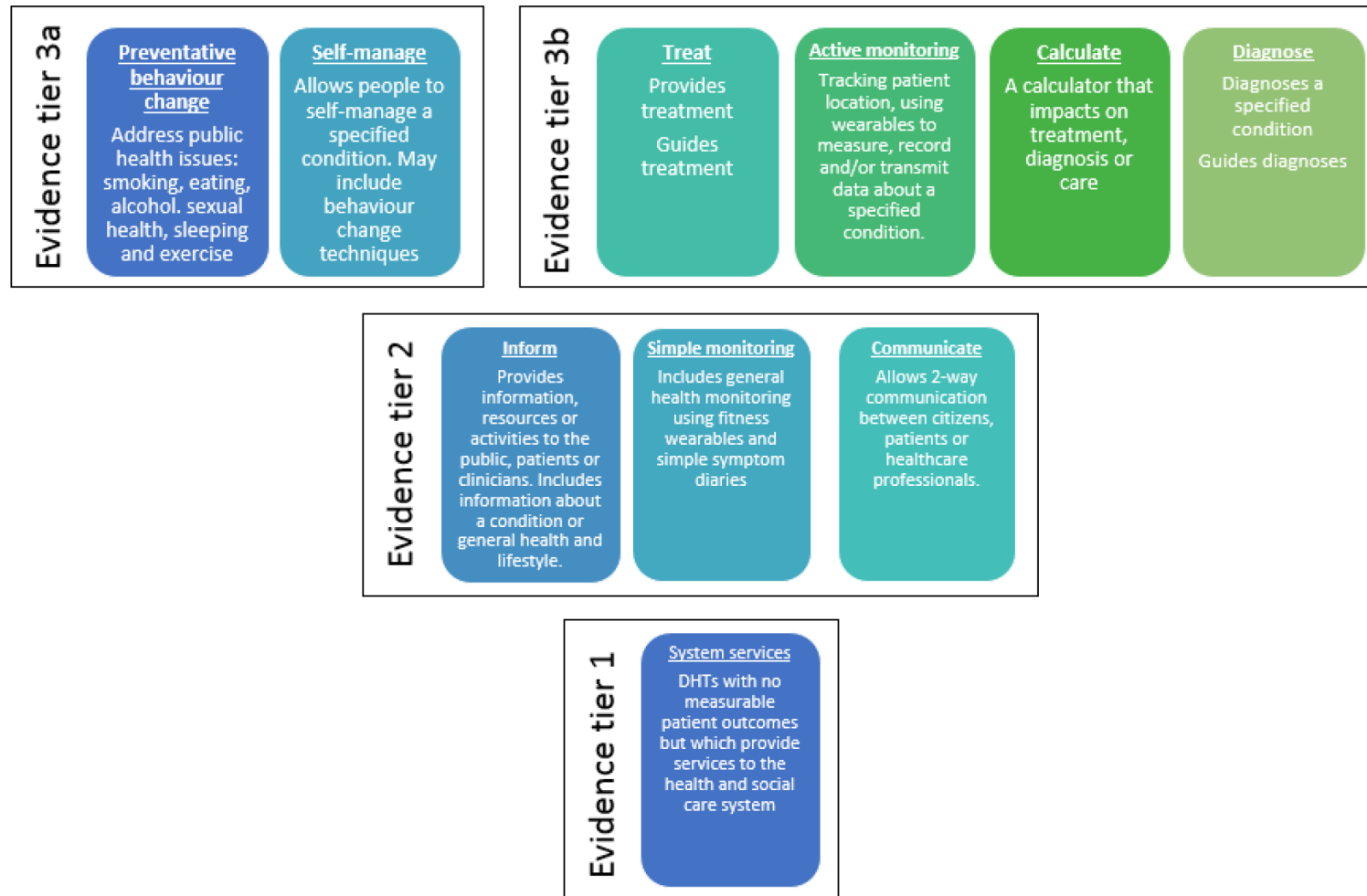


Figure 7-1: DHTs classified by function and stratified into evidence tiers [28]

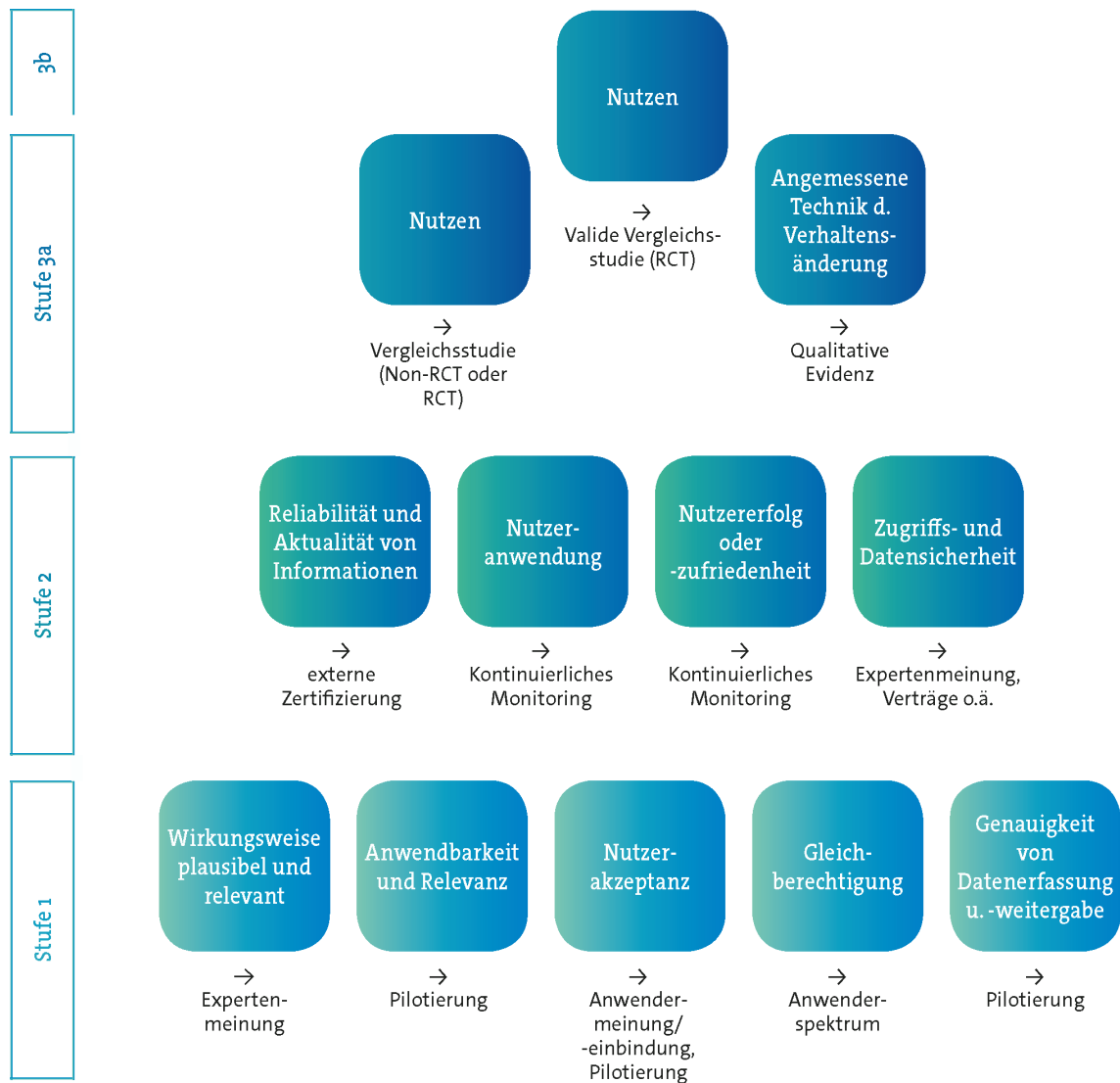


Figure 7-2: NICE Framework Anforderungen an neue digitale Gesundheitsanwendungen, Visualisierung nach IQWiG [29]

7.1.3 Digi-HTA Framework

Table 7-2: HTA domains and criteria of the Digi-HTA framework [30]

HTA-Domain and Criteria	HTA-Domain and Criteria
Company information	Cost
Contact information of company. What is the company's business model? Are quality management systems in use? Which ones?	What are the costs of using the product for a healthcare customer? If the use of the product is free, what is the source of the company's income? What kind of initial costs (estimated minimum and maximum values in detail) does the introduction of the product impose on the organization, including changes to buildings or facilities, a need for new devices and software, as well as needed training? What are the maintenance costs (estimated minimum and maximum values) to the organization for the use of the product? How often must devices or software versions related to the product be renewed? Which uncertainties apply to these cost estimates?
Product information	Effectiveness
The name of the product. Short description of the product. What is the product's readiness level (TRL levels 1–9)? Which platforms and platform versions of the product are available? Does the product have CE and/or FDA approval? Is the product a medical device, and what classification does it have? Is the product classified according to MDD or MDR requirements? Does the product meet the electrical safety requirements for medical devices (if applicable)? Does the use of the product require registration or login? Does the use of the product require strong identification? Does the company have any plans for post-market surveillance of the product? What kind of product support does the company offer? What is the intended use of the product? What are the intended user groups? What problem in the healthcare system is the product trying to solve? Is the aim of the product to replace any existing healthcare services? Does the introduction of the product cause any changes to the premises, information systems, or care processes? Is the product already in use elsewhere in Finland or worldwide? Where, and for how long? What kind of support does the end user need to use the product? If users need training, who organizes it? When? What is the language of training? Does the company have instructions (e.g., a project plan) for healthcare service providers to ensure fluent introduction of the product?	Does the product provide clinical benefits? What are they? Does the product provide benefits to the end users by improving their behavior related to their own health? How so? Does the product provide benefits to the organization (like improving care processes)? How so? What kind of evidence is available for effectiveness (case studies, randomized controlled trials, Cochrane reviews, etc.)? Are there any ongoing studies to investigate the product's effectiveness? Does any institution like the Duodecim Current Care Guidelines recommend the use of the product?
Technical stability	Clinical safety
What is the company's testing process? What is the company's process for handling error messages? Does the company have the capacity to roll back to previous versions of the product? Does the company have a process to proactively monitor the running of systems and system components to automatically identify faults and technical issues? Does the company have a plan for decommissioning the product? Has there been any downtime or impairment time in the use of the product during the last six months?	Are there any risks, possible side effects, or other undesirable effects associated with using the product? Is there any research evidence available related to clinical safety? Have any product-related adverse events been reported or identified? What is the company's process to handle adverse events? Has the product undergone a risk analysis? Are there any undesirable effects associated with misuse of the product? Are the error conditions of guidelines removed, or is their realization unlikely? Is the company aware of the product register and Manufacturer Incident Report supervised by the National Supervisory Authority of Welfare and Health? Who is the responsible person in the company for handling Manufacturer Incident Reports?
Usability and accessibility	Data security and protection
Have all user groups been taken into account in product design, like people with visual or hearing impairments? Has the product been tested with real user groups? What kind of accessibility testing has been performed on the product? Has the functionality of the product been tested with screen readers or other assistive technologies? How have the product's users been taken into account in the product's text (clear, concrete language; the avoidance of professional language)? How have the product's users been taken into account in the design of its textual content (headings, lists, and images)?	Detailed criteria are defined in the following documents: <i>Data Security and Protection Preliminary Task Information</i> <i>Security and Data Protection Requirements</i>
	Artificial intelligence
	Exactly what defined problem is going to be solved by the AI? What is the classification of AI? Visualization only, AI-assisted (e.g., diagnosis/classification/decision), or solely AI-controlled? Could the problem be solved without the AI solution? Is the solution based on machine learning or a neural network? Do the staff have sufficient capacity to understand the operational logic of AI (e.g., do they need additional training)? Are the conclusions and decisions of the AI solution transparent, i.e., can medical staff understand what the decisions are based on? Is the AI solution validated in the environment in which it will be used? What are the data sources for the AI solution?

HTA-Domain and Criteria	HTA-Domain and Criteria
How does the company continue to collect feedback from users and make changes to the product based on this feedback? What changes have been made to the product based on user feedback? How is the company going to continue to evaluate and develop accessibility? Is the product compatible with the following usability guidelines (if applicable)? ■ WCAG 2.0/WCAG 2.1 ■ Papunet Design Guide for Websites EN 301 549 section 11-Software ■ Design guidelines for native application ■ Design guidelines for progressive web application Does the application support OS accessibility features?	Are the data sources used in the training of AI solutions relevant to a final use case (e.g. are the age and gender composition of training groups comparable to that of real user groups)? Are the access rights required for the use of the data in order, and have data protection (e.g., GDPR) and security issues been taken into account? When it comes to classifier teaching, are there enough data relative to the size of the smallest class? Can the AI solution use incomplete data? Can the AI solution use noisy data? Is retraining possible for the AI solution? What are the data sources for retraining? How is it ensured that the system is not taught with irrelevant data? How many tests or results are needed for the AI model? Is the algorithm purchased software as a service (SaaS) or its own design? What performance criteria are used? Does the AI solution change care processes? How? When does the AI solution propose an action? How, and who will actually implement it? Is staff's approval needed for action proposed by the AI?
Interoperability	Robotics
Does the product have interfaces into the website or other software? Does the product have interfaces into the following healthcare services? ■ Electronic patient records (which ones?) Finnish Kanta PHR ■ Other (what?) Are proprietary formats used to store and transfer data? Are the definitions of the original proprietary formats openly available? Does the product have interfaces for other companies' services? Can the data contained in the product be exported in a commonly used or standard format? Does the product use data from other systems via interfaces? If yes, can the data produced by others be separated in the system? Does the product connect with health or wellness devices? If yes, is it compatible with ISO/IEEE 11073 Personal Health Data (PHD) Standards?	Is there any possibility that using the robot could create safety risks for healthcare personnel or customers (e.g., forces that could be destructive or collision with people)? How have those risks been avoided in the robot's design? What kind of arrangements are needed to teach or program the robot to operate? If the robot is battery-operated, what are the operating, idle, and charging times?

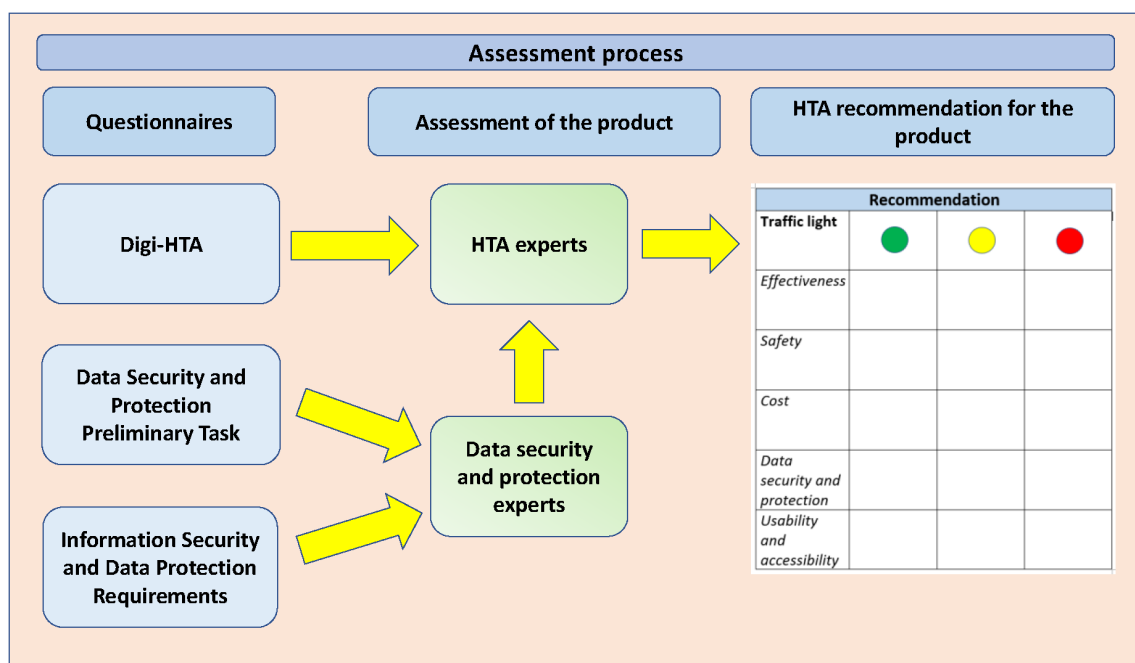


Figure 7-3: Assessment process for digital healthcare services [30]

7.1.4 Digital Health Interventions (DEDHI) Framework

Table 7-3: Overview of the DEDHI framework [32]

1. Preparation Phase
Goals and tasks: To define the conceptual and technological foundation of a DHI. That is, (a) <i>to review existing justificatory knowledge</i> for the development of a DHI (e. g., intervention need, best practices, meta reviews, market analysis), (b) <i>to develop a conceptual model</i> of the DHI outlining the outcomes and intervention components, (c) <i>to conduct a feasibility and acceptability study</i> to test novel DHI components, and finally, (d) <i>to identify an optimization criterion</i> that provides the best expected outcome within technical and health economic constraints.
Technical maturity: Research prototype that provides basic functionality to assess the feasibility of the DHI
Evaluation criteria: Ease of use, adherence, personalization, safety, privacy and security
Implementation barriers: Social interaction, individual characteristics of end user, usability (e. g. user-centred design), expectations (privacy, confidentiality, security), negative associations, workforce, planning, funding (no funding), cost (start-up), standards, regulatory issues
2. Optimization Phase
Goals and tasks: To build an optimized DHI by selecting effective intervention components. That is, (a) <i>to conduct optimization trials</i> (e. g., a micro-randomized trial), and (b) <i>to identify the best DHI configuration that meets the optimization criterion</i> (e. g., a hybrid DHI that applies a chatbot that delivers health literacy information and a human lifestyle coach via a tele-medical online session). If the intervention components are not effective or the optimization criterion is not met, it is suggested to revisit the preparation phase and to make appropriate adaptations.
Technical maturity: Elaborated research prototype that provides the full functionality according to the conceptual model to assess the health impact of DHI components
Evaluation criteria: Effectiveness (individual components of the DHI), perceived benefit, content quality, personalization, perceived enjoyment, aesthetics, adherence, service quality, safety, privacy and security
Implementation barriers: Social support, outcome expectations, usability, funding (for equipment), cost (of technology), integration
3. Evaluation Phase
Goals and tasks: To confirm the effectiveness of an optimized DHI. That is, <i>to conduct a randomized controlled trial</i> to compare the optimized DHI with a control condition (e. g., treatment as usual). If the optimized DHI outperforms the control condition, then the optimized DHI may be implemented at large-scale. If this is not the case, it is suggested to revisit the preparation phase and to make appropriate adaptations.
Technical maturity: Elaborated research prototype that provides the full functionality according to the conceptual model to assess the effectiveness of the DHI
Evaluation criteria: Effectiveness, perceived benefit, adherence, personalization, service quality, safety, privacy and security, accountability
Implementation barriers: Funding (no funding), cost (maintenance), guidelines, methodology
4. Implementation Phase
Goals and tasks: To implement and maintain the effective DHI. That is, (a) <i>to develop a DHI product</i> that can be implemented at large-scale in the healthcare market, (b) <i>to monitor reach, impact and side effects</i> to assess the long-term outreach and efficiency by performing a continuous outreach, impact and side effect assessment, and finally, (c) <i>to update the DHI</i> with respect to content (e. g., if new evidence is available; note: if the update represents a significant change of the causal chain in the conceptual model then it is recommended to reiterate through the preceding phases) and technology (e. g., if application programming interfaces of mobile operating systems are updated or extended, or if security threats make a technical update of the DHI necessary).
Technical maturity: Elaborated research prototype that provides the full functionality according to the conceptual model to assess the health impact of individual intervention components
Evaluation criteria: Adherence, personalization, perceived benefit, content quality, ethics, service quality, safety, privacy and security, accountability
Implementation barriers: Individual resources of end users, expectations (missing accuracy of provided information), usability of technology (technical support, implementation problems), interoperability, human technical support, regional infrastructure, individual characteristics of healthcare providers, negative associations, accessibility, reimbursement, funding (long-term), cost (maintenance), culture

7.1.5 Digital Health Scorecard

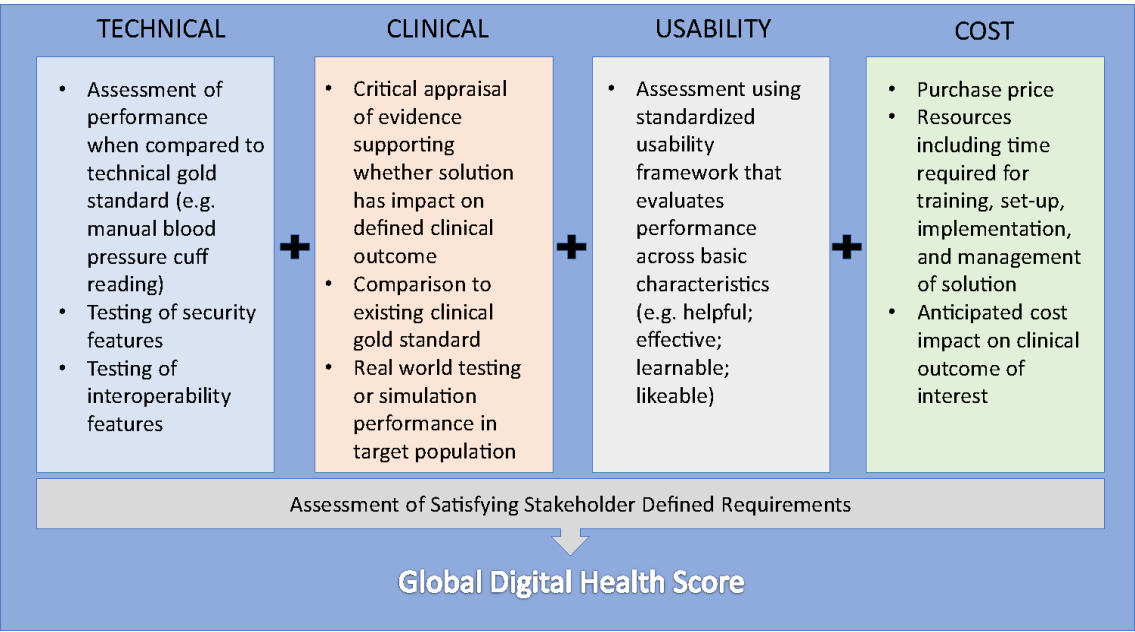


Figure 7-4: Components of Digital Health Scorecard. The four domains of a digital health scorecard with example considerations are detailed in this figure [33]

7.1.6 mHealth Agile Development & Lifecycle

mHealth Agile Development & Evaluation Lifecycle

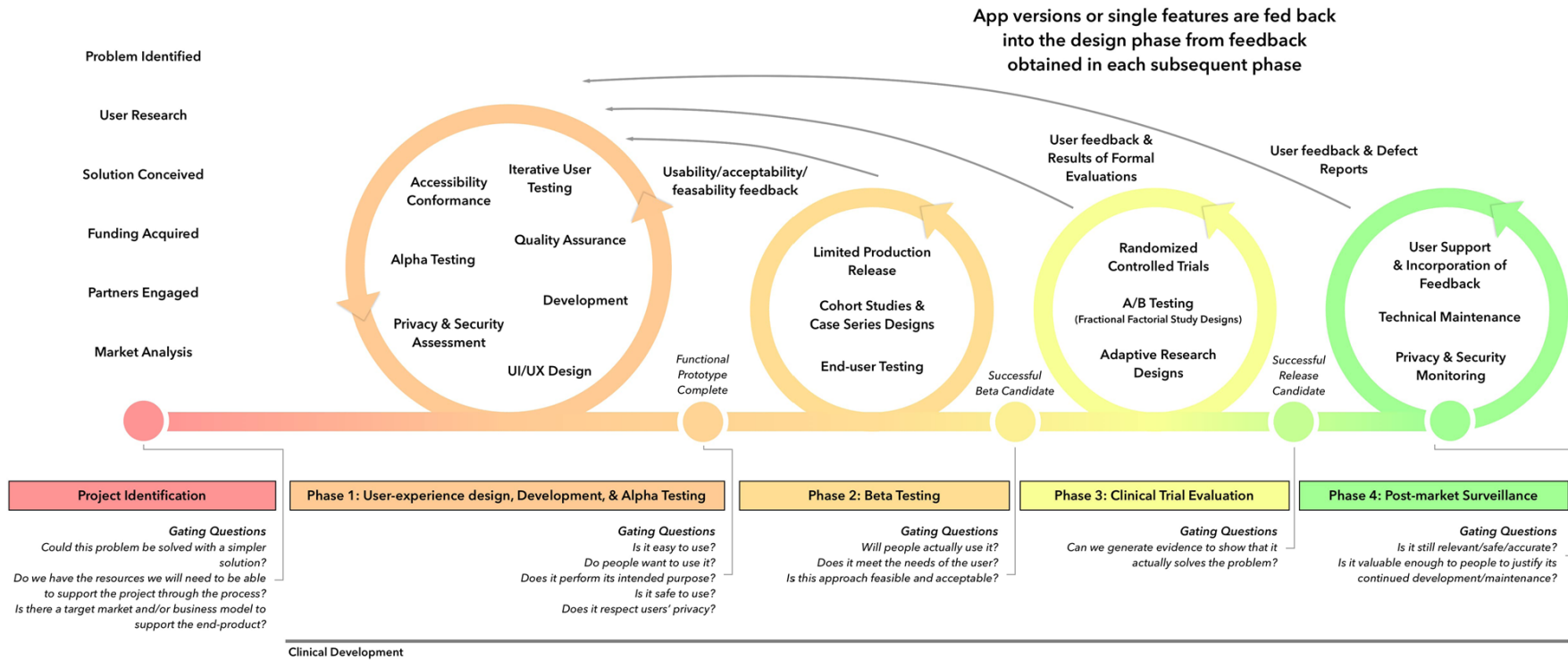


Figure 7-5: mHealth Agile Development & Evaluation Lifecycle [34]

7.2 Extraktionstabellen

7.2.1 INAHTA ListServ Antworten

Table 7-4: Aussagen internationaler HTA-Institutionen zu Assessment Frameworks für digitale Gesundheitsanwendungen

Land	HTA-Institution	Welche Frameworks eignen sich aus Ihrer Sicht in HTA für die Bewertung von DiGA?	Entwicklung/Adaptierung eines neuen/bereits bestehenden Frameworks?	Kommentar
Schottland	HIS	- kein länderspezifisches Framework vorhanden - bisher Verwendung der üblichen HTA Methodik für digitale Technologien (z. B. für die Evaluierung von home health monitoring) aber noch keine durchgeführte Bewertung von DiGA - geplante Pilot Evaluierung anhand des NICE Frameworks	n.a.	Wir wären sehr daran interessiert zu erfahren, wie andere HTA Institutionen vorgehen.
Schweiz	SFOPH	nicht bekannt	nein	
Australien	AHTA	bisher noch nicht	Neu entwickeltes Modul zur Evaluierung mobiler medizinischer Anwendungen (MMA) als Ergänzung zum allgemeinen HTA-Prozesse (by Moshi et al.)	Modul wurde noch nicht verwendet, da noch kein MMA zur Refundierung eingereicht wurde. Modul wird allerdings bei der Überarbeitung der Richtlinien für die Bewertung von Refunderungsanträgen an das Medical Services Advisory Committee (MSAC) referenziert werden (vorauss. Fertigstellung in Oktober).
UK	HTW	NICE Evidence Standards Framework (für die Auswahl von Technologien + während des Bewertungsprozesses)	Nein. Nach Konsultation mit Interessensvertretern, ob die NICE Standards auch für Wales anwendbar sind, wurden keine Bedenken geäußert.	
Spanien	AQuAS	Hauptsächlich wird das MAST framework verwendet, welches auf dem EUnetHTA core model basiert. NICE Framework?	- Veröffentlichung eines Frameworks für die Bewertung von mHealth Anwendungen in 2017 (Bradway et al. 2017) Hauptbestandteile: - Gerechte Verteilung von Verantwortlichkeiten, Involvement aller Institutionen, die einen relevanten Beitrag leisten - Summative (d. h. während oder nach der Implementierung) und formative (während der gesamten Entwicklungs) Evaluation 2 Phasen der Bewertung: - Pre-assessment von mHealth Anwendungen - Checkliste (länderspezifisch, sollte u. a. Verwendungszweck der Intervention, Stand der Entwicklung, Sicherheit und Privatsphäre etc. enthalten) - Aktuell wird ein Instrument zur Beurteilung von mHealth Produkten entwickelt (vorauss. in Q3 2020 auf Spanisch verfügbar)	

Land	HTA-Institution	Welche Frameworks eignen sich aus Ihrer Sicht in HTA für die Bewertung von DiGA?	Entwicklung/Adaptierung eines neuen/bereits bestehenden Frameworks?	Kommentar
Uruguay		HAD	Bis jetzt wurde noch keine mHealth Applikation bewertet	
Kanada	CADTH	Verwendung des standard deliberative HTA Framework für DiGA aber relevante Aspekte aus anderen Frameworks (z. B. NICE) sollen bei der Bewertung auch berücksichtigt werden.	Bisher noch nicht. Zusammenarbeit mit der Regulierungsbehörde (Health Canada).	Health Canada besitzt ebenfalls kein eigenes Framework. Health Canada hat zwar eine Orientierungshilfe zum Thema Software als Medizinprodukt erstellt, aber noch keine zu umfassenderen Guidance zu DiGA gegeben.
	INESSS	n.a.	Entwicklung eines multidimensionalen Frameworks, mit dem eine Vielzahl von Aspekten erfasst werden sollen, die allgemein für HTA und eben auch für DiGA von Relevanz sind. Das Framework baut auf dem Accountability for reasonableness (A4R) ethical framework auf und beruht dabei auf reflexiven multikriteriellen Ansätzen. Die Möglichkeit, die klinischen, wirtschaftlichen, organisatorischen, bevölkerungsbezogenen und soziopolitischen Dimensionen von DiGA unter der Berücksichtigung von ethischen Aspekten zu betrachten, ermöglicht es, potenzielle positive und negative Auswirkungen der evaluierten Technologien auszugleichen, sodass gerechte Empfehlungen abgegeben werden können.	
Deutschland	IQWiG	Das IQWiG empfiehlt, das NICE Framework anzuwenden. Im Hinblick auf ein mögliches Rahmenwerk von Evidenzanforderungen sollten für digitale wie für nicht-digitale medizinische Interventionen, etwa therapeutische oder diagnostische, dieselben Standards angewendet werden. Für die Gesundheit ist weniger Evidenz als für medizinische Interventionen erforderlich. Nicht-medizinische digitale Gesundheitsanwendungen sollen die Gesundheit verbessern, umfassen aber keine medizinische Versorgung oder greifen direkt in diese ein.	nein	2 Wege für die Erstattung von DiGA in Deutschland: 1. Produkte mit höherem Risiko (CE-gekennzeichnet als Klasse IIb oder III) erfordern ein vollständiges HTA wie jede andere medizinische Intervention. In diesem Fall sind der G-BA und das IQWiG für die Evidenzbeurteilung zuständig. 2. Neues Verfahren seit 2020 für digitale Produkte mit geringem Risiko (Medizinprodukte der Klasse I oder IIb) das geringere Evidenzanforderungen stellt. In der Verordnung des Bundesgesundheitsministeriums wird beschrieben, welche Daten für einen Antrag als neue DiGA beim BfArM eingereicht werden müssen. Nach Ansicht des IQWiG sollte nicht allein die Risikoklasse eines Medizinprodukts für die Klassifizierung von DiGA verantwortlich sein, sondern App-Funktionen müssten ebenfalls Berücksichtigung finden.
	G-BA	Für digitale Apps gibt es keine spezifischen Frameworks sie werden nach den gleichen Prinzipien wie andere Verfahren bewertet.	n.a.	In Deutschland werden DiGA Gesundheitsanwendungen in Abhängigkeit von der Risikoklasse bewertet. Für die Risikoklassen I und IIa liegt die Verantwortung beim BfArM. Anträge, die in die Risikoklasse IIb oder III eingestuft werden, werden vom G-BA bewertet, sofern sie Teil eines Verfahrens sind und wenn ein Antrag beim G-BA eingereicht wurde.

Land	HTA-Institution	Welche Frameworks eignen sich aus Ihrer Sicht in HTA für die Bewertung von DiGA?	Entwicklung/Adaptierung eines neuen/bereits bestehenden Frameworks?	Kommentar
Schweden	SBU	Es werden die gleichen Methoden und Frameworks wie für andere Interventionen verwendet. Folglich ist die Qualitätsbeurteilung vom Studiendesign abhängig.	Bis jetzt nicht, aber interessiert daran von anderen Ländern zu lernen	Das Thema mobile Apps wurde in einem „SBU Policy Support“ Bericht zum Thema Patientenbeteiligung basierend auf einer Übersichtsarbeit von Whitehead L, Seaton P. betrachtet. Der SBU-Bericht ist nur auf Schwedisch verfügbar. (SBU Policy Support fasst wissenschaftliche Evidenz für die Politikberatung, Entscheidungsfindung sowie die Entwicklung nationaler Richtlinien zusammen.) In Schweden gibt es eine eHealth Agency, welche sich ausschließlich mit e-Health befasst aber bisher noch keine Bewertungen vornimmt.
Ukraine	SEC	n.a.	nein	
Italien	UVT	■ Das in Moshi et al. 2020 (Tabelle 1) vorgeschlagene Vorgehen stimmt am meisten mit dem italien. Ansatz überein. ■ Es benötigt ein erweitertes Kernmodell, einen multi-domainen Ansatz, um die Besonderheiten der verschiedenen Arten von DiGA zu erfassen. Digitale und organisatorische Aspekte sollten detaillierter untersucht werden. ■ Verweis auf die im HAS Good Practice Guidelines on Health Apps and Smart Devices aufgeführten Kriterien (d. h. Kriterien in Bezug auf technisches Design, Datenfluss, generierte Inhalte, Standardisierung usw.) ■ Weitere zu berücksichtigende Elemente: Entwicklungsstand der Technologie, die verfügbare Evidenz, die Validierung klinischer Endpunkte, die Integration in die aktuellen klinischen Versorgungspfade, die erforderlichen Ressourcen und den bereitgestellten Zusatznutzen, Koordination/Integration der verschiedenen digitalen Gesundheitstechnologien sowie mittel- und langfristige Nachhaltigkeit der Technologie	In einer Fallstudie wurde ein erweitertes Kernmodell anhand einer mobilen Gesundheitsanwendung getestet. Ziel war es, alle zusätzlichen HTA-Elemente, die für die Bewertung dieser Art von Technologie relevant sind, zu identifizieren sowie Wissen und Kompetenz für die Evaluierung dieser neuen Art von Technologie zu erwerben. Unsere Bewertung war Teil eines Projekts zur Einführung und Implementierung einer DiGA in die klinische Praxis.	
Frankreich	HAS	n.a.	n.a.	Vorschlag eines Erfahrungsaustausches/Webinars im Herbst zum Thema digital health. Dies wäre eine gute Vorbereitung für das F2F-Treffen, bei dem das Thema ebenfalls aufgeführt wird. Wir könnten uns abstimmen, wann wir unsere nächste TC zum Thema INAHTA F2F haben werden.
Brasilien	CONITEC	Bisher hat Conitec noch nicht erörtert, welches Framework sich für den Einsatz in HTA von DiGA für Brasilien eignet.	Zurzeit entwickelt oder adaptiert Conitec keine Frameworks zur Bewertung mobiler Gesundheitsanwendungen.	

7.3 Suchstrategien

Suchstrategie für Cochrane Datenbank

Search date: 10/06/2020	
ID	Search
#1	MeSH descriptor: [Mobile Applications] explode all trees
#2	(mHealth) (Word variations have been searched)
#3	(m-Health) (Word variations have been searched)
#4	(„mobile Health“) (Word variations have been searched)
#5	(eHealth) (Word variations have been searched)
#6	(e-Health) (Word variations have been searched)
#7	(„digital health“) (Word variations have been searched)
#8	(DigiHealth) (Word variations have been searched)
#9	(Digi-Health) (Word variations have been searched)
#10	(„health app“) (Word variations have been searched)
#11	(„health apps“) (Word variations have been searched)
#12	(healthapp*) (Word variations have been searched)
#13	(„medical app“):ti,ab,kw (Word variations have been searched)
#14	(„medical apps“) (Word variations have been searched)
#15	(„mobile app“) (Word variations have been searched)
#16	(„mobile apps“) (Word variations have been searched)
#17	((mobile OR health OR medical) NEAR application*) (Word variations have been searched)
#18	(„standalone software*“) (Word variations have been searched)
#19	(„closed loop system*“) (Word variations have been searched)
#20	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 (Word variations have been searched)
#21	(assess* OR eval* OR apprais* OR HTA* OR framework* OR reimburs* OR re-imburs* OR „decision support*“):ti,ab,kw (Word variations have been searched)
#22	MeSH descriptor: [Insurance, Health, Reimbursement] explode all trees
#23	MeSH descriptor: [Decision Support Systems, Clinical] explode all trees
#24	MeSH descriptor: [Technology Assessment, Biomedical] explode all trees
#25	#21 OR #22 OR #23 OR #24 (Word variations have been searched)
#26	#20 AND #25 with Cochrane Library publication date Between Mar 2018 and Jun 2020, in Cochrane Reviews, Cochrane Protocols, Clinical Answers, Editorials, Special collections (Word variations have been searched)
Total: 88 Hits	

Suchstrategie für CRD

#### mHealth Frameworks (RJ/CW) 170620	
Search date: 17.06.2020	
1	MeSH DESCRIPTOR Mobile Applications EXPLODE ALL TREES
2	(mHealth)
3	(m-Health)
4	(mobile Health)
5	(eHealth)
6	(e-Health)
7	(digital health)
8	(DigiHealth)

9	(Digi-Health)
10	(health app*)
11	(medical app*)
12	(mobile app*)
13	(standalone software*)
14	(stand-alone software*)
15	(closed loop system*)
16	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15
17	(#16) WHERE LPD FROM 01/03/2018 TO 17/06/2020
0 Hits	

Suchstrategie für Embase

No.	Query Results	Results	Date
#41.	#40 AND [1-3-2018]/sd NOT [11-6-2020]/sd	526	10 Jun 2020
#40.	#34 OR #39	1,019	10 Jun 2020
#39.	#33 AND #38	121	10 Jun 2020
#38.	#35 OR #36 OR #37	25,305	10 Jun 2020
#37.	hta*:ti,ab,kw,de	8,563	10 Jun 2020
#36.	'technology assessment':ti,ab,kw,de	19,990	10 Jun 2020
#35.	'biomedical technology assessment'/exp	14,574	10 Jun 2020
#34.	#33 AND 'Review'/it	917	10 Jun 2020
#33.	#25 AND #32	9,924	10 Jun 2020
#32.	#26 OR #27 OR #28 OR #29 OR #30 OR #31	8,117,833	10 Jun 2020
#31.	'appraisal of guidelines for research and evaluation ii'/exp	18	10 Jun 2020
#30.	'appraisal'/exp	23	10 Jun 2020
#29.	'framework'/exp	12	10 Jun 2020
#28.	'decision support system'/exp	25,039	10 Jun 2020
#27.	'reimbursement'/exp	56,298	10 Jun 2020
#26.	assess*:ti,ab OR eval*:ti,ab OR apprais*:ti,ab OR hta*:ti,ab OR framework*:ti,ab OR reimburs*:ti,ab OR 're-imburs*:ti,ab OR 'decision support*:ti,ab	8,082,423	10 Jun 2020
#25.	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #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	19,672	10 Jun 2020
#24.	'closed loop system*':ti,ab	2,173	10 Jun 2020
#23.	'closed loop system'/exp	73	10 Jun 2020
#22.	'standalone software*':ti,ab	125	10 Jun 2020
#21.	(mobile NEAR/1 application*):ti,ab	3,312	10 Jun 2020
#20.	'mobil app*':ti,ab	1	10 Jun 2020
#19.	mobilapp*:ti,ab	-	10 Jun 2020
#18.	'mobile apps':ti,ab	997	10 Jun 2020
#17.	'mobile app':ti,ab	1,680	10 Jun 2020
#16.	'medical apps':ti,ab	235	10 Jun 2020
#15.	'medical app':ti,ab	68	10 Jun 2020
#14.	'healthapp*':ti,ab	3	10 Jun 2020
#13.	'health apps':ti,ab	512	10 Jun 2020
#12.	'health app':ti,ab	255	10 Jun 2020
#11.	'digi health':ti,ab	-	10 Jun 2020
#10.	digihealth:ti,ab	-	10 Jun 2020
#9.	'digital health':ti,ab	1,568	10 Jun 2020

#8.	'digital health'/exp	48	10 Jun 2020
#7.	'e health':ti,ab	3,205	10 Jun 2020
#6.	ehealth:ti,ab	2,845	10 Jun 2020
#5.	'm-health':ti,ab	554	10 Jun 2020
#4.	mhealth:ti,ab	2,864	10 Jun 2020
#3.	'mhealth'/exp	33	10 Jun 2020
#2.	'mobile health':ti,ab	3,498	10 Jun 2020
#1.	'mobile health application'/exp	885	10 Jun 2020

Suchstrategie für Medline

Database: Ovid MEDLINE(R) and In-Process & Other Non-Indexed Citations and Daily <1946 to June 09, 2020>, Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily <2016 to June 09, 2020>	
Search Strategy:	
1	exp Telemedicine/(37169)
2	exp Mobile Applications/(10149)
3	1 and 2 (2441)
4	mHealth.ti,ab. (5293)
5	m-Health.ti,ab. (624)
6	mobile Health.ti,ab. (5947)
7	eHealth.ti,ab. (4090)
8	e-Health.ti,ab. (2889)
9	digital health.ti,ab. (2455)
10	DigiHealth.ti,ab. (0)
11	Digi-Health.ti,ab. (0)
12	health app\$1.ti,ab. (1300)
13	healthapp*.ti,ab. (2)
14	medical app\$1.ti,ab. (324)
15	mobile app\$1.ti,ab. (3623)
16	(mobile adj2 application*).ti,ab. (6040)
17	standalone software.ti,ab. (196)
18	closed loop system*.ti,ab. (2502)
19	3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 (27871)
20	(assess* or eval* or apprais* or HTA* or framework* or reimburs* or re-imburs* or decision support*).ti,ab. (7747057)
21	exp Insurance, Health, Reimbursement/(49700)
22	exp Decision Support Systems, Clinical/(9924)
23	20 or 21 or 22 (7783378)
24	19 and 23 (14227)
25	review*.pt. (3232180)
26	technology assessment*.mp. (17929)
27	HTA*.ti,ab. (6051)
28	25 or 26 or 27 (3248708)
29	24 and 28 (1922)
30	limit 29 to (english or german) (1889)
31	limit 30 to dt=20180301-20200610 (877)
32	remove duplicates from 31 (450)
Search date: 10/06/2020	

Suchstrategie für inaHTA

Query Nr.	Search query, "Hits", "Searched At"
20	(closed loop system*) OR (stand-alone software*) OR (stand alone software*) OR (standalone software*) OR (mobile app*) OR (medical apps) OR (medical app) OR (healthapp*) OR (health apps) OR (health app) OR (Digi Health) OR (DigiHealth) OR (digital health) OR (eHealth) OR (mobile Health) OR (mHealth) OR (Telemedicine"[mhe]) OR („Mobile Applications"[mhe])", "188", "limited to years 2018-2020
19	(closed loop system*) OR (stand-alone software*) OR (stand alone software*) OR (standalone software*) OR (mobile app*) OR (medical apps) OR (medical app) OR (healthapp*) OR (health apps) OR (health app) OR (Digi Health) OR (DigiHealth) OR (digital health) OR (eHealth) OR (mobile Health) OR (mHealth) OR (Telemedicine"[mhe]) OR („Mobile Applications"[mhe])", "188", "2020-06-17T14:34:29.000000Z"
18	closed loop system*, "3", "2020-06-17T14:34:02.000000Z"
17	stand-alone software*, "6", "2020-06-17T14:33:22.000000Z"
16	stand alone software*, "0", "2020-06-17T14:33:07.000000Z"
15	standalone software*, "0", "2020-06-17T14:32:54.000000Z"
14	mobile app*, "21", "2020-06-17T14:31:59.000000Z"
13	medical apps, "0", "2020-06-17T14:31:16.000000Z"
12	medical app, "0", "2020-06-17T14:31:07.000000Z"
11	healthapp*, "0", "2020-06-17T14:30:51.000000Z"
10	health apps, "5", "2020-06-17T14:30:23.000000Z"
9	health app, "0", "2020-06-17T14:30:06.000000Z"
8	Digi Health, "0", "2020-06-17T14:29:26.000000Z"
7	DigiHealth, "0", "2020-06-17T14:29:03.000000Z"
6	digital health, "47", "2020-06-17T14:28:42.000000Z"
5	eHealth, "1", "2020-06-17T14:27:13.000000Z"
4	mobile Health, "30", "2020-06-17T14:26:42.000000Z"
3	mHealth, "0", "2020-06-17T14:26:17.000000Z"
2	Telemedicine"[mhe]", "106", "2020-06-17T14:25:39.000000Z"
1	Mobile Applications"[mhe]", "0", "2020-06-17T14:21:36.000000Z"
Date of Search: 17.06.2020	
4 Hits	



HTA Austria

Austrian Institute for
Health Technology Assessment
GmbH