

RESEARCH

Open Access



# Evidence requirements of permanently listed digital health applications (DiGA) and their implementation in the German DiGA directory: an analysis

Melanie Mäder<sup>1,2\*</sup>, Patrick Timpel<sup>2</sup>, Tonio Schönfelder<sup>2,3</sup>, Carsta Militzer-Horstmann<sup>1,2</sup>, Sandy Scheibe<sup>2,3</sup>, Ria Heinrich<sup>2</sup> and Dennis Häckl<sup>1,2</sup>

## Abstract

**Background** With its digital health application (DiGA)-system, Germany is considered one of Europe's pioneers in the field of evidence-based digital health. Incorporating DiGA into standard medical care must be based on evidence-based success factors; however, a comprehensive overview of the evidence required of scientific studies for their approval is lacking.

**Objective** The study aims to, (1) identify specific requirements defined by the Federal Institute for Drugs and Medical Devices (German: Bundesinstitut für Arzneimittel- und Medizinprodukte; BfArM) to design adequate studies, proving a positive healthcare effect, and (2) to assess the evidence given for applications permanently listed in the DiGA directory.

**Methods** A multi-step approach was used: (1) identification of the evidence requirements for applications permanently listed in the DiGA directory, (2) identification of the evidence available supporting them.

**Results** All DiGA permanently listed in the DiGA directory (13 applications) are included in the formal analysis. Most DiGA addressed mental health ( $n = 7$ ), and can be prescribed for one or two indications ( $n = 10$ ). All permanently listed DiGA have demonstrated their positive healthcare effect through a medical benefit, and most of them provide evidence for one defined primary endpoint. All DiGA manufacturers conducted a randomized controlled trial.

**Discussion** It is striking that— although patient-relevant structural and procedural improvements show high potential for improving care, especially in terms of processes — all DiGA have provided a positive care effect via a medical benefit. Although BfArM accepts study designs with a lower level of evidence for the proof of a positive healthcare effect, all manufacturers conducted a study with a high level of evidence.

**Conclusion** The results of this analysis indicate that permanently listed DiGA meet higher standards than required by the guideline.

**Keywords** Digital health applications, DiHA, DiGA, Digital health technology, Evidence evaluation, Evaluation concept

\*Correspondence:

Melanie Mäder

melanie.maeder@wig2.de

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Digital health applications (known by their German abbreviation, DiGA) are increasingly present, with health systems around the world creating different legal frameworks for their integration into standard care. Together with Belgium, Germany was one of the first countries in Europe to develop an official framework to reimburse DiGA use. Other countries, such as England, Denmark, the Netherlands, Norway, Sweden, and the USA, are still in the development process [1]. France has already announced its intention to adopt the German system [2].

DiGA are certified medical devices with a primarily digital function, through which the medical purpose is achieved. Mainly used by patients, they detect, monitor, treat, or alleviate disease, injury, or disability. Shared use between patients and healthcare providers is also possible, however devices that merely read and transmit data are not considered DiGA [3]. They offer the potential of improved health outcomes [4], higher health standards, and improved and equal access to health services [5]. In addition, DiGA cost-effectively improve patient care [6].

The Federal Republic of Germany's national parliament passed The Act to Improve Healthcare Provision through Digitalization and Innovation (Digital Healthcare Act – DVG) in 2019. With this law, healthcare providers were given the option of prescribing DiGA. To be part of standard care, the DiGA must be listed in the directory in accordance with paragraph 139e of the fifth book of social code (§ 139e SGB V), first requiring evidence-based proof of a benefit. The benefit is defined by the manufacturer guideline of the Federal Institute for Drugs and Medical Devices (German: Bundesinstitut für Arzneimittel- und Medizinprodukte; BfArM) [3]. First, certification as a medical device in a low-risk class (I or IIa) according to the Medical Device Regulation (MDR) is required. Basic requirements are set for data protection, interoperability, robustness, and user-friendliness, but the central criterion is proof that the DiGA has a positive healthcare effect. DiGA manufacturers must provide this evidence with a comparative scientific study in an International Classification of Diseases (ICD-10)-defined patient group. The positive healthcare effect can be a medical benefit and/or a patient-relevant structural and procedural improvement. The fast-track process offers two paths into the DiGA directory; if all criteria are fulfilled and BfArM reaches a positive decision, the DiGA is permanently included in the DiGA directory; if no positive healthcare effect has yet been demonstrated, the DiGA can be provisional included. In case of the latter, the expected positive healthcare effect must be proven within a period of 12 months with a previously approved evaluation study [3].

The DiGA manufacturer guideline specifies what type of evidence from positive healthcare effect studies is acceptable [3, 7]. After using the fast-track DiGA approval process, Brönneke et al. (2021) highlighted the importance of measures by legislation to help decide if innovations benefit patients in standard healthcare [7]. Similarly, Heimann et al. (2021) described factors required using the fast-track listing, including internal or self-commissioned external audits before submitting the application, consultation with BfArM on positive healthcare effects, and responses to queries from BfArM [8]. Löbker et al. (2021) reported on their experience from consultations on DiGA showing most (80%) of DiGA directory application manufacturers have taken advantage of a consultation during the process. The rate of withdrawn/rejected applications was higher, if manufacturers had not sought advice (63%), compared to manufacturers who had previously discussed key content with BfArM (35%) [9]. Düvel et al. (2021) qualitatively identified potential solutions to improve DiGA access to statutory standard care, recommending a central advisory office [10]. Lantzsch et al. (2022) concluded that there is room for improvement in the fast-track process, particularly in study reporting, as well as in outcomes for patient-relevant improvement of structure and processes. They demand new study designs to pave the way for the use of real-world data [11]. To generate high-quality evidence of DiGA positive healthcare effects, Stern et al. (2022) recommends further research on the impact of missing data, study endpoints, control group, multimodal interventions, study question, equity, generalizability, confounders and fit for purpose [12]. Geier et al. (2021) described different perspectives of the German Digital Health Association on DiGA, arguing the need for accompanying research to address the specific challenges of study design and methods in generating evidence for DiGA [13]. Hemkens et al. (2021) also emphasized the relevance of a robust evidence-based benefit assessment in DiGA approvals, stating that sustainable and efficient DiGA benefit assessment requires continuously adjusted evaluation in everyday care; central to this are randomized study designs that are integrated into standard care [14].

Incorporating DiGA into standard medical care must be based on evidence-based success factors; however, a comprehensive overview of the evidence required of scientific studies for their approval is lacking [15]. We want to investigate which methodologies used to generate evidence-based proof of benefit have been successfully implemented, based on the already permanently listed DiGA.

Our study aims to, (1) identify specific requirements defined by BfArM to design adequate studies, proving a

positive healthcare effect, and (2) to assess the evidence given for applications permanently listed in the DiGA directory.

**Methods**

All permanently listed applications in the German DiGA directory (as of November 15<sup>th</sup>, 2022) were included in this study.

The study used a multi-step approach: First, we identified the evidence requirements for applications permanently listed in the DiGA directory, based on an analysis of the BfArM manufacturer guideline (as of March 18<sup>th</sup>, 2022) using the PICOS scheme: Population, Intervention, Control, Outcome, Study Design [16]. Then, a single researcher extracted these requirements, and then they were double-checked by a second independent researcher. All extracted data were transferred and categorized to a data extraction sheet (MS Excel). Finally, the evidence available supporting the permanently listed DiGA was identified, using the following sources: DiGA directory; study registries (German registry of clinical studies (DRKS), clinical trials.gov, ISRCTN registry); published study protocols; published study reports; submitted publications of permanently listed DiGA; and finally, manufacturer websites.

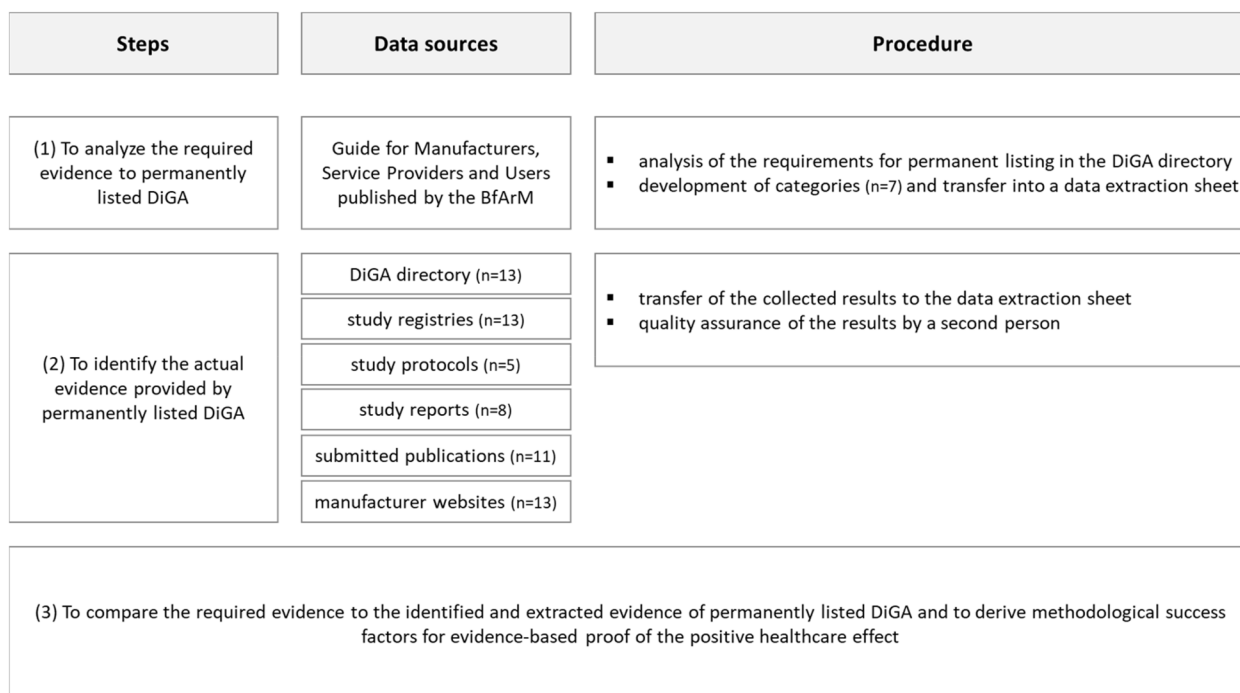
The sources (including studies, study protocols, reports) were used to extract data using the pre-defined

data extraction sheet. Data extraction was performed by one researcher, followed by quality assurance of extracted information carried out by a second researcher. The methodological approach is represented in Fig. 1.

**Results**

All DiGA permanently listed in the DiGA directory (13 applications) are included in the formal analysis (Table 1). Most of these were initially accepted into the DiGA directory with an application directly for permanent listing. Four DiGA manufacturers (Kalmeda, Selfapy Depression, Vivira and Zanadio) initially applied for a provisional listing during testing, before being permanently listed. All four applications with initial provisional inclusion in the DiGA directory extended the trial period; Kalmeda by three months, Selfapy Depression and Vivira by four months, and Zanadio by ten months.

We identified seven relevant categories from our BfArM manufacturer guideline document analysis: (1) patient population, (2) positive healthcare effect and study endpoints, (3) study design, (4) study location, (5) observation period and observation times, (6) sample size and drop-out, and (7) study results. These categories are used to guide the step-by-step reporting of findings.



**Fig. 1** Methodological approach

BfArM = Federal Institute for Drugs and Medical Devices (German: Bundesinstitut für Arzneimittel- und Medizinprodukte, DiGA = Digital health application

**Table 1** Overview of indication area and quantity of ICD-10-Codes (three, four and five-digit)

DiGA	Indication area	Quantity ICD-10 3 digit	Quantity ICD-10 4 digit	Quantity ICD-10 5 digit
Deprexis [17, 18]	mental health	0	6	0
Elevida [19]	nervous system	1	0	0
Hello Better Diabetes and Depression [20, 21]	hormones and metabolism	2	0	0
Hello Better Panik [22, 23]	mental health	0	1	1
Hello Better Stress and Burnout [24, 25]	mental health	1	0	0
Hello Better Vaginismus Plus [26, 27]	mental health	0	2	0
Kalmeda [28]	ears	0	1	0
Selfapy Depression [29]	mental health	0	2 (9) <sup>a</sup>	0
Somnio [30]	mental health	0	1	0
Velibra [31]	mental health	0	3	1
Vivira [32]	muscles, bones, joints	0	11 (16) <sup>a</sup>	9 (29) <sup>a</sup>
Vorvida [33]	nervous system	0	2	0
Zanadio [34]	hormones and metabolism	0	0	2

DiGA Digital health application, ICD International Classification of Diseases

<sup>a</sup> Originally targeted number of indications

### Patient population and patient characteristics

A positive healthcare effect must be provided for at least one defined patient population. The delineation is based on one or more indications listed in the ICD-10 catalog; this international classification provides a specific disease definition. Three- or four-digit ICD-10 codes are permissible. This can be used to distinguish whether the positive healthcare effect should be demonstrated for all patients with a specific disorder (e.g., F32.-Depressive Episode) or for a specific patient population (e.g., F32.0-Mild Depressive Episode). If several indications are given, the evidence must be provided for each defined patient group. Thus, the ICD-10 code chosen is a measure of the specificity and concreteness of the patient population addressed. If the positive healthcare effect among the indications is comparable, the proof can be provided for several indications combined [3].

Of 13 DiGA applications, seven address indications in mental health. One DiGA works at the interface of mental health and metabolism (Hello Better Diabetes and Depression), and the others are relevant to the nervous system (Elevida, Vorvida), the musculoskeletal system (Vivira), ears (Kalmeda), and hormones and metabolism (Zanadio). The majority ( $n=10$ ) of the listed DiGA can be prescribed for one or two indications, except for Deprexis with six, Velibra with four and Vivira with 20 indications. The manufacturers of Selfapy Depression originally sought listing for nine indications but were only able to prove a positive healthcare effect for two of these. Vivira manufacturers were able to provide evidence of a positive healthcare effect for 20 of the 45 indications they originally sought listing for. Table 1 shows the DiGA

indication areas addressed, and the number of ICD-10 codes (separated into three-, four-, and five-digit codes) for which listing in the DiGA directory was achieved.

Except for Hello Better Vaginismus Plus, all permanently listed DiGA were developed for both women and men. In all studies, the proportion of female participants was higher than that of men. This trend was evident in both intervention groups (IG) and control groups (CG). Zanadio was approved for women only, due to an insufficient number of males in the trial. All DiGA manufacturers defined a minimum age of 18 years as inclusion criterion. On average, the study participants' mean age was between 28.0 (Hello Better Vaginismus Plus) and 51.3 (Hello Better Diabetes and Depression) years. In all studies already published (except for Hello Better Vaginismus Plus) the average age of the participants was between 35 and 55 years. There were no significant differences in age between IG and CG. To date, participants with a higher education and a permanent job are more frequently represented in DiGA studies (Table 2).

### Positive healthcare effect and study endpoints

For permanent or provisional inclusion in the DiGA directory, manufacturers must demonstrate one or more positive healthcare effects by means of a scientific study. This proof can be provided as a medical benefit (improvement of the state of health, reduction of the duration of the disease, prolongation of survival, or improvement in the quality of life), and/or a patient-relevant improvement of structure and processes (e.g., coordination of treatment processes, alignment of treatment with guidelines and recognized standards, or adherence).

**Table 2** Overview of patient population and patient characteristics

DiGA	Patient population		Patient characteristics					
	Gender	Age	Gender		Age		Education and/or employment status	
			IG	CG	IG	CG	IG	CG
<b>Deprexis</b>								
Study 1 [17]	m / f	18-65 years	f=58 (74.4 %) m=20 (25.6 %)	f=64 (75.3 %) m=21 (24.7 %)	M=44 (11.02)	M=40 (11.48)	<u>Educational status:</u> Lower second-ary=4 (5.1 %) Middle second-ary=23 (29.5 %) Higher second-ary=14 (17.9 %) Highest second-ary=37 (47.4 %)	<u>Educational status:</u> Lower second-ary=3 (3.5 %) Middle second-ary=14 (16.5 %) Higher second-ary=12 (14.1 %) Highest second-ary=53 (62.4 %)
Study 2 [18]	m / f	18-65 years	f=350 (68.8 %) m=159 (31.2 %)	f=345 (68.5 %) m=159 (31.5 %)	M=42.8 (11.0)	M=42.9 (11.0)	<u>Educational status:</u> Not yet completed=2 (0.4 %) No degree=1 (0.2 %) Lower second-ary school=29 (5.7 %) Middle second-ary school=131 (25.8 %) Higher second-ary school qualifying for university of applied sci-ence=87 (17.1 %) Higher second-ary school qualifying for university=249 (48.9 %) Other=10 (2.0 %)	<u>Educational status:</u> Not yet completed=0 (0.0 %) No degree=0 (0.0 %) Lower secondary school=24 (4.8 %) Middle second-ary school=112 (22.2%) Higher secondary school qualifying for university of applied sci-ence=85 (16.8 %) Higher secondary school qualify-ing for univer-sity=271 (53.8%) Other=12 (2.4 %)
							<u>Employment status:</u> Full time=38 (48.7 %) Regular part-time=9 (11.5 %) Not working=31 (39.7 %)	<u>Employment status:</u> Full time=31 (36.5 %) Regular part-time=16 (18.8 %) Not working=38 (44.7 %)
							<u>Employment status:</u> Full time=220 (43.3 %) Part time=117 (23.0 %)	<u>Employment status:</u> Full time=214 (42.4 %) Part time=114 (22.6 %)

**Table 2** (continued)

DiGA	Patient population		Patient characteristics					
	Gender	Age	Gender		Age		Education and/or employment status	
			IG	CG	IG	CG	IG	CG
<b>Elevida</b> [19]	m / f	≥ 18 years	f=114 (82.0 %) m=25 (18 %)	f=108 (79 %) m=28 (21 %)	M=40.8 (11.1)	M=41.9 (9.4)	<u>Employment status:</u> Full-time/part-time=71 (51 %) Housemaker/student/other=13 (9 %) Unemployed/unable to work/retired= 55 (40 %)	<u>Employment status:</u> Full-time/part-time=64 (47 %) Housemaker/student/other=11 (8 %) Unemployed/unable to work/retired=61 (45 %)
<b>Hello Better Diabetes and Depression</b> [20, 21]	m / f	≥ 18 years	f=81 (63.3%) m=47 (36.7 %)	f=79 (62.7%) m=47 (37.3 %)	M=50.2 (11.6)	M=51.3 (11.9)	<u>Educational status:</u> Low=19 (14.8 %) Middle=76 (59.4 %) High=32 (25.0 %) <u>Employment status:</u> Employed=77 (60.2 %)	<u>Educational status:</u> Low=17 (13.5 %) Middle=64 (50.8 %) High=45 (35.7 %) <u>Employment status:</u> Employed=79 (62.7 %)
<b>Hello Better Panik</b> [22, 23]	m / f	> 18 years	f=27 (60 %) m=18 (40 %)	f=24 (51.1 %) m=23 (48.9 %)	M=39.33 (10.83)	M=37.43 (10.03)	<u>Educational status:</u> Low=3 (6.7 %) Middle=11 (24.4 %) High=31 (68.9 %) <u>Employment status:</u> Full-time=27 (60.0 %) Part-time=15 (33.3 %) Not employed=3 (6.7 %) Unemployed=0 Currently sick=0	<u>Educational status:</u> Low=4 (8.5 %) Middle=14 (29.8 %) High=29 (61.7 %) <u>Employment status:</u> Full-time=25 (53.2 %) Part-time=15 (31.9 %) Not employed=3 (6.4 %) Unemployed=1 (2.1 %) Currently sick=3 (6.4 %)
<b>Hello Better Stress and Burnout</b> [24, 25]	m / f	≥ 18 years	f=97 (73.5 %) m=35 (26.5 %)	f=96 (72.7 %) m=36 (27.3 %)	M=42.4 (10.7)	M=44.2 (9.6)	<u>Educational status:</u> Low=3 (2.3 %) Middle=25 (18.9 %) High=104 (78.8 %) <u>Employment status:</u> Full-time=105 (79.5 %) Part-time=25 (18.9 %) Currently sick=2 (1.5 %)	<u>Educational status:</u> Low=2 (1.5 %) Middle=31 (23.5 %) High=99 (75.0 %) <u>Employment status:</u> Full-time=99 (75.0 %) Part-time=32 (24.2 %) Currently sick=1 (0.8 %)

**Table 2** (continued)

DiGA	Patient population		Patient characteristics					
	Gender	Age	Gender		Age		Education and/or employment status	
			IG	CG	IG	CG	IG	CG
<b>Hello Better Vaginismus Plus</b> [26, 27]	f	≥ 18 years	f=100 (100 %) m=0	f=100 (100 %) m=0	M=29.46 (9.82)	M=28.04 (7.84)	<u>Educational status:</u> Low=0 (0.00 %) Middle=42 (42.00 %) High=58 (58.00 %)	<u>Educational status:</u> Low=4 (4.00 %) Middle=47 (47.00 %) High=49 (49.00 %)
<b>Kalmeda</b> [28]	m / f	≥ 18 years	f= 49 (52.1%) m=45 (47.9%)	f=41 (44.1%) m=52 (55.9%)	M=48.1 (12.8)	M= 48.4 (12.2)	n.r.	n.r.
<b>Selfapy Depression</b> [29]	m / f	18-65 years	IG1: f=126 (83.4 %) m=25 (16.6 %) IG2: f=126 (84.0 %) m=24 (16.0 %)	f=81 (81.0 %) m=19 (19.0 %)	IG1: M=38 (10.7) IG2: M=37 (10.8)	M=36 (11.9)	<u>Employment status:</u> Employee: IG 1=82 (54.3 %), IG2=86 (57.3 %) Self-employed: IG1=3 (2.0 %), IG2=4 (2.7 %) Trainee: IG1=12 (7.9 %), IG2=6 (4.0 %) Other: IG1=7 (4.6 %), IG2=3 (2.0 %) Not reported: IG1=47 (31.3 %), IG2=51 (34.0 %)	<u>Employment status:</u> Employee=57 (57.0 %) Self-employed=2 (2.0 %) Trainee=25 (25.0 %) Other=14 (14.0 %) Not reported=2 (2.0 %)
<b>Somnio</b> [30]	m / f	≥ 18 years	f=21 (72 %) m=8 (28 %)	f=18 (67 %) m=9 (33 %)	M=41.72 (17.31)	M=44.04 (20.05)	<u>Educational status:</u> Obligatory school=1 (3 %) Basic apprenticeship=3 (10 %) Higher apprenticeship=6 (21 %) College=5 (17 %) University=11 (38 %) Other=3 (10 %)	<u>Educational status:</u> Obligatory school=2 (7 %) Basic apprenticeship=5 (19 %) Higher apprenticeship=1 (4 %) College=4 (15 %) University=11 (41 %) Other=4 (15 %)
							<u>Employment status:</u> Self-employed=3 (10 %) Employee=17 (59 %) Retired=4 (14 %) Housewife/ man=1 (3 %) Unemployed=2 (7 %) Other=2 (7 %)	<u>Employment status:</u> Self-employed=3 (11 %) Employee=13 (48 %) Retired=4 (15 %) Housewife/ man=2 (7 %) Unemployed=3 (11 %) Other=2 (7 %)



**Table 2** (continued)

DiGA	Patient population		Patient characteristics					
	Gender	Age	Gender		Age		Education and/or employment status	
			IG	CG	IG	CG	IG	CG
<b>Velibra [31]</b>	m / f	≥ 18 years	f=48 (69 %) m=22 (31 %)	f=50 (72 %) m=19 (28 %)	M=42.1 (12.2)	M=41.8 (12.2)	<u>Educational status:</u> Compulsory school=5 (7 %) Apprenticeship=32 (46 %) College=9 (13 %) University=24 (34 %)	<u>Educational status:</u> Compulsory school=4 (6 %) Apprenticeship=25 (36 %) College=10 (14 %) University=30 (43 %)
<b>Vivira [32]</b>	m / f	≥ 18 years	<i>No data published yet</i>					
<b>Vorvida [33]</b>	m / f	≥ 18 years	f=170 (56 %) m=136 (44 %)	f=149 (49 %) m=153 (51 %)	M=40.4 (11.2)	M=40.7 (12.1)	<u>Educational status:</u> Apprenticeship/technical school=203 (66.3 %) University/college=78 (25.5 %) No vocational training=17 (5.6 %) Other=8 (2.6 %)	<u>Educational status:</u> Apprenticeship/technical school=185 (61.3 %) University/college=97 (32.1 %) No vocational training=18 (6.0 %) Other=2 (0.7 %)
<b>Zanadio [34]</b>	m / f	18-65 years	f=n.r. m=7	f=n.r. m=6	<i>No data published yet</i>			
							<u>Employment status:</u> Employed=236 (77.1 %) Not employed=45 (14.7 %) Learning=15 (4.9 %) Other=10 (3.3 %)	<u>Employment status:</u> Employed=239 (79.1 %) Not employed=41 (13.6 %) Learning=20 (6.6 %) Other=2 (0.7 %)

DiGA Digital health application, IG Intervention group, CG Control group, m Male, f Female, M (SD) Mean (standard deviation)

The positive healthcare effect to be demonstrated must relate directly to the insured person and be proven within the framework of the study to be conducted by means of defined endpoints [3].

All permanently listed DiGA have demonstrated their positive healthcare effect through a medical benefit. The manufacturers of Velibra and Vorvida demonstrated a patient-relevant structural and procedural improvement,

in addition to the medical benefit. The most frequently proven medical benefit of all permanently listed DiGA can be categorized as improvement in the state of health. The Vivira application manufacturers additionally aimed for a reduced disease duration and an improved quality of life, but were unable to provide evidence for either. The Velibra manufacturers demonstrated a patient-relevant improvement of structure and process, reducing



therapy-related efforts of patients and their relatives, and the Vorvida manufacturers, an improvement in patient autonomy (Table 3). Most DiGA provide evidence for one defined primary endpoint, except the Velibra manufacturers, who defined a total of five primary endpoints and proved four of them (the subscale, mental health, of the Short Form Health Survey-12, could not be proven). Depending on the defined patient group, various primary endpoints are attempted. The operationalization of the primary endpoints was carried out using validated questionnaires (Table 5).

### Study design

A quantitative comparative (retrospective or prospective) study is required to provide evidence of a positive healthcare effect. Studies intended to prove a positive healthcare effect should reflect the actual healthcare reality and be conducted with data close to the healthcare system. A comparison can be made intra-individually or inter-individually. In an intra-individual, single-arm comparison, the data can, for example, be collected before and after the application of the DiGA. In an inter-individual, two-arm comparison, the intervention group data (DiGA use) are compared with control group data. The following options exist for the design of the control group: (1) treatment without the use of a DiGA, (2) non-treatment, and (3) treatment with another, comparable DiGA already permanently listed in the DiGA directory at the time of application. It is important that the selection of the control group corresponds to the standard of care [3].

All permanently listed DiGA manufacturers conducted a randomized controlled trial (RCT) using an inter-individual comparison of an intervention group with a control group. Almost all DiGA were tested against standard of care. The manufacturers of Selfapy Depression originally designed two intervention groups: one with, and one without psychological support. Evidence could only be proved for DiGA use without psychological support. Most control groups (standard of care) were offered access to the DiGA after the end of the study. Three DiGA manufacturers provided additional materials to the participants: Hello Better Diabetes and Depression enabled the control group to use an online knowledge transfer program, Kalmeda provided general information on the topic of tinnitus, and the manufacturers of Selfapy Depression sent weekly emails with standardized mindfulness exercises (Table 3).

### Study location

Studies to prove a positive healthcare effect must be conducted in Germany, as the actual healthcare setting is closely linked to the positive healthcare effects. The comparison of the intervention against a control group

without DiGA use is only meaningful, if a treatment in the German healthcare system is addressed. If the comparability of the healthcare situation can be proven, a study in other countries is also permissible [3].

Evidence of positive healthcare effect was obtained for 11 DiGA in Germany, and for two applications in Switzerland (Table 3).

### Observation period and observation times

There are no concrete specifications of observation periods and times in the BfArM guideline. Baseline data collection is considered reasonable; the data collection periods (incl. possible follow-ups after the intervention phase) should be described [3].

The observation periods ranged from six weeks (Somnio) to nine months (Kalmeda, Zanadio). Most DiGA studies had an observation period of eight to twelve weeks. In addition to collecting data at baseline (t0) and at the fixed primary endpoint collection time, ten DiGA studies also collected follow-up data. In most cases, follow-up data were collected after six and/or twelve months. Almost all studies date back several years, so that existing data was used to prove a positive healthcare effect (Table 3).

### Sample size and drop-out

BfArM requires studies to have an adequate sample size calculation/planning. In confirmatory studies, the sample sizes should be estimated based on the primary outcome measure and the relevant effect size. BfArM also requires studies to record the number of drop-outs and the respective reasons for these, with a transparent presentation in a study flow chart. A high drop-out rate should also be considered in connection with the respective clinical picture. Studies on addictive diseases, for example, show higher drop-out rates in actual care reality compared to diseases with recognized therapies and high responder rates. A high drop-out rate does not indicate the success of the application and can only be accepted to a limited extent [3].

The manufacturers of Somnio were able to provide evidence of a positive healthcare effect with the lowest sample size ( $n=56$ ), and the manufacturers of Deprexis used the largest sample ( $n=1,013$ ). Proof was provided for Hello Better Panik based on 92 cases. Five studies had sample sizes between 100 and 199, and six studies between 200 and 299. All other studies worked with sample sizes larger than 300. The distribution of the individuals to the intervention and control groups was done by means of 1:1 randomization in most studies; one exception was the Selfapy Depression study, with two intervention groups (IG1=151, IG2=150, CG=100). The drop-out rates at the time of the primary endpoint survey

**Table 3** Overview of positive healthcare effects; indications, study designs, observation durations, sample sizes, IG, CG

DiGA	Positive care effect	Medical benefit	Patient-relevant improvement of structure and processes	Study design	Observation duration and survey times	Sample size	IG	CG	Location of study	Time period of study <sup>b</sup>
<b>Deprexis</b> Study 1 [17]	medical benefit	improvement in the state of health	n.a	RCT	<b>3 months</b> t1 = 3 months t2 = 6 months follow-up	<b>n = 163</b> IG = 78 CG = 85	DiGA + UC UC + waiting group	UC + waiting group	Hamburg, Lübeck, Bern	02/2013–05/2014
Study 2 [18]						<b>n = 1013</b> IG = 509 CG = 504			Berlin, Bielefeld, Hamburg, Lübeck	07/2012–02/2017
<b>Elevida</b> [19]	medical benefit	improvement in the state of health	n.a	RCT	<b>12 weeks</b> t1 = 12 weeks t2 = 24 weeks follow-up	<b>n = 275</b> IG = 139 CG = 136	DiGA + UC	UC + waiting group	Hamburg	07/2014–05/2016
<b>Hello Better Diabetes and Depression</b> [20, 21]	medical benefit	improvement in the state of health	n.a	RCT	<b>8 weeks</b> t1 = 8 weeks t2 = 6 months follow-up 1 t3 = 12 months follow-up 2	<b>n = 254</b> IG = 128 CG = 126	DiGA + UC	UC + waiting group + online program for knowledge transfer	Lüneburg	03/2013–01/2015
<b>Hello Better Panik</b> [22, 23]	medical benefit	improvement in the state of health	n.a	RCT	<b>8 weeks</b> t1 = 8 weeks t2 = 3 months follow-up 1 t3 = 6 months follow-up 2	<b>n = 92</b> IG = 45 CG = 47	DiGA + UC	UC + waiting group	Lüneburg	08/2013–05/2015
<b>Hello Better Stress and Burnout</b> [24, 25]	medical benefit	improvement in the state of health	n.a	RCT	<b>7 weeks</b> t1 = 7 weeks t2 = 6 months follow-up 1 t3 = 12 months follow-up 2, only IG	<b>n = 264</b> IG = 132 CG = 132	DiGA + UC	UC + waiting group	Lüneburg	03/2013–11/2014
<b>Hello Better Vaginismus Plus</b> [26, 27]	medical benefit	improvement in the state of health	n.a	RCT	<b>12 weeks</b> t1 = 12 weeks t2 = 6 months follow-up	<b>n = 200</b> IG = 100 CG = 100	DiGA + UC	UC + waiting group	Erlangen	04/2016–9/2018
<b>Kalmeda</b> [28]	medical benefit	improvement in the state of health	n.a	RCT	<b>9 months (IG)</b> <b>12 months (CG)</b> t1 = 3 months (IG) t2 = 9 months (IG) t3 = 12 months (CG)	<b>n = 187</b> IG = 94 CG = 93	DiGA	waiting group + general information about tinnitus	Duisburg	08/2020–04/2022

**Table 3** (continued)

DIGA	Positive care effect	Medical benefit	Patient-relevant improvement of structure and processes	Study design	Observation duration and survey times	Sample size	IG	CG	Location of study	Time period of study <sup>b</sup>
<b>Selfapy Depression [29]</b>	medical benefit	improvement in the state of health	n.a	RCT	<b>12 weeks</b> t1 = 6 weeks t2 = 12 weeks t3 = 24 weeks follow-up	<b>n = 401</b> IG1 = 151 IG2 = 150 CG = 100	IG1: DiGA with psychological support (proof not provided) IG2: DiGA without any support	waiting group + weekly standardized mindfulness exercises via email	Berlin	08/2019–n.r
<b>Sommio [30]</b>	medical benefit	improvement in the state of health	n.a	RCT	<b>6 weeks</b> t1 = 6 weeks t2 = 12 months follow-up	<b>n = 56</b> IG = 29 CG = 27	DiGA	waiting group	Zurich	01/2016–10/2016
<b>Velibra [31]</b>	medical benefit + patient-relevant improvement of structure and processes	improvement in the state of health	reduction of therapy-related efforts and strains for patients and their relatives	RCT	<b>9 weeks</b> t1 = 9 weeks t2 = 6 months follow-up	<b>n = 139</b> IG = 70 CG = 69	DiGA + UC	UC + waiting group	Bern	05/2014–05/2016
<b>Vivira [32]</b>	medical benefit	improvement in the state of health	n.a	RCT	<b>12 weeks</b> t1 = 2 weeks t2 = 6 weeks t3 = 12 weeks	<b>n = 215</b>	DiGA	UC (physiotherapy)	Tübingen	11/2020–03/2021
<b>Vorvida [33]</b>	medical benefit + patient-relevant improvement of structure and processes	improvement in the state of health	patient autonomy	RCT	<b>6 months</b> t1 = 3 months t2 = 6 months	hip = 256 knee = 402 <b>n = 608</b> IG = 306 CG = 302	DiGA + UC	UC + waiting group	Hamburg	03/2015–ongoing
<b>Zanadio [34]</b>	medical benefit	improvement in the state of health	n.a	RCT	<b>9 months</b> t1 = 3 months t2 = 6 months t3 = 9 months t4 = 12 months follow-up	<b>n = 149</b>	DiGA	UC	online in Germany	02/2021–ongoing

DIGA Digital health application, IG Intervention group, CG Control group, RCT Randomized controlled trial, UC Usual care, n.a. Not applicable

ranged from 6.7% (Zanadio) to 30.0% (Vorvida). In the pre-post study of Vivira (intended to transfer the results to other indications), drop-out rates of more than 90% were observed, which is why the evidence was assessed as not provided. Drop-out rates were higher in intervention groups than in control groups, except for two studies (for Hello Better Panik and Selfapy Depression applications). Drop-out rates at the last follow-up time point ranged from 5.9% (Somnio, a respondent who missed t1 provided data at t2) to 61.3% (Selfapy Depression) (Table 4).

### Study results

A positive healthcare effect is considered proven if the outcome is (clinically) relevant, patient-relevant, and statistically significant [3].

All studies were able to deliver significant differences between the intervention and control groups at the primary endpoint and thus, significant results. The effect sizes, according to Cohen et al. were small effect sizes (<0.5) for three DiGA; medium effect sizes (0.5–0.8) for six applications, and in the range of large effect sizes (>0.8) for four DiGA. No effect sizes were reported for

four DiGA. For the Deprexis study, the effect sizes of the two published studies were used, and for the Velibra study, the effect sizes were those related to the four primary endpoints. The effect sizes increased for five DiGA at the follow-up time point, and decreased for two DiGA. No effect sizes at the follow-up time point were reported for six DiGA. DiGA that reported the minimum clinically relevant difference also exceeded this effect size (Table 5).

After analyzing the requirements and implementation of the evidence of permanently listed DiGA, our research question, which methodological success factors regarding evidence-based proof of benefit can be derived from the already permanently listed DiGA, can be summarized with the following Fig. 2.

The requirements stated in the BfArM guideline only provide a frame of reference. The factors identified in the present analysis can help successfully implement these requirements in the standard of care.

### Further results

Although this was not a study objective, we have found the DiGA directory to lack reporting quality and

**Table 4** Overview of sample sizes and drop-outs

DiGA	Case number			Drop-out at primary survey time point			Total drop-out at the end of the study		
	total	IG	CG	total	IG	CG	total	IG	CG
<b>Deprexis</b>									
Study 1 [17]	163	78	85	29 (17.8%)	17	12	45 (27.6%)	21	24
Study 2 [18]	1013	509	504	219 (21.6%)	114	105	259 (25.6%)	131	128
<b>Elevida</b> [19]	275	139	136	51 (18.5%)	36	15	64 (23.3%)	44	20
<b>Hello Better Diabetes and Depression</b> [20, 21]	254	128	126	47 (18.5%)	31	16	74 (29.1%)	50	24
<b>Hello Better Panik</b> [22, 23]	92	45	47	8 (8.7%)	4	4	20 (21.7%)	10	10
<b>Hello Better Stress und Burnout</b> [24, 25]	264	132	132	21 (8.0%)	16	5	51 (19.3%)	40	11
<b>Hello Better Vaginismus Plus</b> [26, 27]	200	100	100	30 (15.0%)	22	8	51 (25.5%)	42	9
<b>Kalmeda</b> [28]	187	94	93	24 (12.8%)	16	8	92 (48.9%)	41	51
<b>Selfapy Depression</b> [29]	401	IG1 = 151 IG2 = 150	100	95 (23.7%)	IG1 = 19 IG2 = 30	46	246 (61.3%)	IG1 = 87 IG2 = 88	71
<b>Somnio</b> [30]	56	29	27	4 (7.1%)	4	0	3 (5.4%)	3	0
<b>Velibra</b> [31]	139	70	69	19 (13.7%)	13	6	26 (37.0%)	26	n.a
<b>Vivira</b> [32]									
RCT	215	2	n.r	n.r	n.r	n.r	n.r	n.r	n.r
Pre-Post-Study	hip = 256 knee = 402	n.r	n.r	hip = 242 (94.5%) knee = 372 (92.5%)	n.r	n.r	n.r	n.r	n.r
<b>Vorvida</b> [33]	608	306	302	183 (30.0%)	114	69	183 (30.1%)	114	69
<b>Zanadio</b> [34]	149	n.r	n.r	10 (6.7%)	n.r	n.r	n.r	n.r	n.r

DiGA = Digital health application, IG = Intervention group, CG = Control group, n.r. = Not reported, n.a. = Not applicable

**Table 5** Overview of primary endpoints and study results (based on the intention-to-treat analyses)

DiGA	Positive care effect	Primary endpoint	Operationalization	Interpretation	MCID	Results
<b>Deprexis</b>						
Study 1 [17]	improvement of the state of health	depression severity	Patient Health Questionnaire-9 (PHQ-9)	0–27 higher scores indicate stronger symptoms level of depression: 0–4 = minimal 5–9 = mild 10–14 = moderate 15–19 = moderately severe 20–27 = severe [35]	5-point reduction [35]	<p>t0 (Baseline) IG: M = 16.62 (3.44) CG: M = 17.20 (3.86)</p> <p>t1 (3 months)<sup>a</sup> IG: M = 10.08 (6.37) CG: M = 13.64 (6.14) <b>p &lt; 0.01</b> Cohens d = 0.57 [0.22, 0.92]</p> <p>t2 (6 months follow-up) IG: M = 11.28 (6.04) CG: M = 13.39 (6.59) Cohens d = 0.33 [0.3, 0.69]</p>
Study 2 [18]						<p>t0 (Baseline) IG: M = 10.23 (2.41) CG: M = 10.34 (2.40)</p> <p>t1 (12 weeks)<sup>a</sup> IG: M = 7.54 (4.04) CG: M = 9.15 (4.30) Cohens d = 0.39 [0.13, 0.64]</p> <p>t2 (24 weeks follow-up) Intergroup difference -2.19 [-0.57, -3.82] <b>p = 0.0080</b> Cohens d = 0.53</p> <p>t3 (12 months follow-up 2) IG: M = 19.88 (10.15) KG: M = 26.73 (10.43) MD = 5.57 (2.02) <b>p = 0.002</b> Cohens d = 0.63 [0.18, 1.08]</p>
<b>Elevida [19, 36]</b>						
	improvement of the state of health	fatigue	Chalder Fatigue Scale	0–33 higher scores indicate stronger symptoms [37]	between 0.7 and 1.4 [37, 38]	<p>t0 (Baseline) IG: M = 21.58 (5.32) CG: M = 21.17 (5.02)</p> <p>t1 (8 weeks)<sup>a</sup> Intergroup difference -2.74 [-1.16, -4.32] <b>p = 0.0007</b> Cohens d = 0.53</p> <p>t2 (6 months follow-up 1) Intergroup difference -2.19 [-0.57, -3.82] <b>p = 0.0080</b></p>
<b>Hello Better Diabetes and Depression [20, 21]</b>						
	improvement of the state of health	depressive symptom severity	Allgemeine Depressionsskala (ADS)	0–60 higher scores indicate stronger symptoms [39] score of 23 or higher is indicative of clinically relevant depressive symptoms in German populations [40]	standardized mean difference $\geq 0.24$ [41]	<p>t0 (Baseline) IG: M = 32.22 (6.96) CG: M = 31.55 (7.57)</p> <p>t1 (8 weeks)<sup>a</sup> IG: M = 20.86 (9.73) CG: M = 29.1 (9.08) MD = 8.07 (1.37) <b>p &lt; 0.001</b> Cohens d = 0.94 [0.62, 1.25]</p> <p>t2 (6 months follow-up 1) IG: M = 20.09 (11.15) CG: M = 27.13 (10.14) MD = 5.78 (1.95) <b>p &lt; 0.001</b> Cohens d = 0.62 [0.21, 1.04]</p> <p>t3 (12 months follow-up 2) IG: M = 19.88 (10.15) KG: M = 26.73 (10.43) MD = 5.57 (2.02) <b>p = 0.002</b> Cohens d = 0.63 [0.18, 1.08]</p>

**Table 5** (continued)

DiGA	Positive care effect	Primary endpoint	Operationalization	Interpretation	MCID	Results
Hello Better Panik [22, 23]	improvement of the state of health	severity of panic and agoraphobia symptoms	Panic Symptom Severity and Self-Rating (German: PAS)	0–52 higher scores indicate stronger symptoms 0–8 = no clinically relevant symptoms 9–28 = moderate symptoms > 29 = severe level of symptoms [42]		<b>t0 (Baseline)</b> IG: M = 18.18 (6.54) CG: M = 19.43 (5.49)
						<b>t1 (8 weeks)<sup>a</sup></b> IG: M = 11.73 (6.90) CG: M = 17.60 (8.86) MD = 5.99 (1.86) <b>p = 0.009</b> Cohens d = 0.78 [0.31, 1.26]
Hello Better Stress and Burnout [24, 25]	improvement of the state of health	level of perceived stress	Perceived Stress Scale (PSS-10)	0–40 higher scores indicate stronger symptoms 0–13 = low 14–26 = moderate 27–40 = high perceived stress [43]		<b>t0 (Baseline)</b> IG: M = 25.89 (3.85) CG: M = 25.15 (3.96)
						<b>t1 (7 weeks)<sup>a</sup></b> IG: M = 17.88 (6.17) CG: M = 22.96 (6.07) <b>p &lt; 0.001</b> Cohens d = 0.83 [0.58–1.08]
Hello Better Vaginismus Plus [26, 27]	improvement of the state of health	intercourse penetration behavior	Primary Endpoint Questionnaire (PEQ)	0 = not attempted or unsuccessful 1 = attempted and sometimes successful or attempted and always successful [44]		<b>t0 (Baseline)</b> IG: M = 0 CG: M = 0
						<b>t1 (12 weeks)<sup>a</sup></b> IG = 31 (31%) CG = 13 (13%) Chi-Quad-rat = 9.44 <b>p &lt; 0.01</b> OR = 3.01 [1.46–6.18]
						<b>t2 (6 months follow-up 1)</b> IG: M = 9.66 (7.65) CG: M = 16.32 (7.13) MD = 6.72 (1.97) <b>p = 0.009</b> Cohens d = 1.00 [0.42, 1.58]
						<b>t2 (6 months follow-up)</b> IG: M = 16.08 (6.03) CG: M = 22.10 (5.81) <b>p &lt; 0.001</b> Cohens d = 1.02 [0.76–1.27]
						<b>t3 (12 months follow-up 2)</b> IG: M = 8.63 (6.33) CG: M = 14.68 (7.65) MD = 6.07 (1.90) <b>p = 0.009</b> Cohens d = 0.96 [0.37, 1.56]
						<b>t3 (12 months follow-up 2)</b> IG: M = 16.25 (6.35)

**Table 5** (continued)

DiGA	Positive care effect	Primary endpoint	Operationalization	Interpretation	MCID	Results
<b>Kalmeda [28]</b>	improvement of the state of health	tinnitus exposition	Tinnitus questionnaire according to Göbel and Hiller	0–84 higher scores indicate stronger symptoms 0–30 = mild 31–46 = medium 47–59 = severe 60–84 = severely [45]		<p><b>t0 (baseline)</b> IG: M = 37.14 (13.01) CG: M = 37.46 (14.29)</p> <p><b>t1 (3 months)<sup>a</sup></b> IG: M = 23.43 (11.77) CG: M = 24.76 (15.71) <b>p &lt; 0.0001</b></p> <p><b>t2 (9 months) only IG</b> IG: M = 17.61 (10.21) <b>p &lt; 0.0001</b></p> <p><b>t3 (12 months) only CG</b> CG: M = 20.27 (13.01) <b>p &lt; 0.0001</b></p>
<b>Selfapy Depression [29]</b>	improvement of the state of health	depressive symptoms	Beck-Depressions-Inventar II (BDI-II)	0–63 higher scores indicate stronger symptoms 0–8 = no depression 9–13 = minimal depression 14–19 = mild depression 20–28 = moderate depression 29–63 = severe depression [46, 47]		<p><b>t0 (Baseline)</b> IG1: M = 30.09 (9.18) IG2: M = 30.54 (8.53) CG: M = 30.88 (10.74)</p> <p><b>t1 (6 weeks)</b> IG1: M = 20.71 (6.98) IG2: M = 22.51 (7.83) CG: M = 29.09 (6.39) IG1 vs. IG2: Cohens d = 0.24 [0.02, 0.48] IG1 vs. CG: Cohens d = 1.25 [0.99, 1.54] IG2 vs. CG: Cohens d = 0.92 [0.65, 1.20]</p> <p><b>t2 (12 weeks)<sup>a</sup></b> IG1: M = 16.61 (9.55) IG2: M = 18.49 (8.88) CG: M = 30.26 (6.97) IG1 vs. IG2: Cohens d = 0.20 [− 0.04, 0.45] <b>p = 0.18</b> IG1 vs. CG: Cohens d = 1.63 [1.37, 1.93] <b>p &lt; 0.001</b> IG2 vs. CG: Cohens d = 1.47 [1.22, 1.73] <b>p &lt; 0.001</b></p> <p><b>t3 (24 weeks follow-up)</b> for the 24-week follow-up assessment only per protocol data were used</p>



**Table 5** (continued)

DiGA	Positive care effect	Primary endpoint	Operationalization	Interpretation	MCID	Results
<b>Somnio [30]</b>	improvement of the state of health	insomnia	Insomnia severity index (ISI)	0–28 higher scores indicate stronger symptoms 0–7 = normal 8–14 = sub-threshold insomnia 15–21 = moderate insomnia > 21 = severe clinical insomnia [48]		<p><b>t0 (Baseline)</b></p> <p>IG: M = 15.38 (3.74) CG: M = 13.26 (3.29)</p> <p><b>t1 (6 weeks)<sup>a</sup></b></p> <p>IG: M = 7.80 (4.94) CG: M = 12.04 (3.86) t = 5.06 <b>p &lt; 0.001</b> Cohens d = 1.79</p> <p><b>t2 (12 months follow-up)</b></p> <p>IG: M = 7.36 (5.38) Cohens d = 1.98 [1.31, 2.66]</p>
<b>Velibra [31, 49]</b>	improvement of the state of health	depression, anxiety and tension/stress	Depression Anxiety Stress Scales – Short Form (DASS-21)	0–120 higher scores indicate stronger symptoms [50]	$\Delta = 6.42$ [51] d = 0.24 [41]	<p><b>t0 (Baseline)</b></p> <p>IG: M = 58.2 (24.4) CG: M = 55.8 (21.3)</p> <p><b>t1 (9 weeks)<sup>a</sup></b></p> <p>IG: M = 40.9 (25.7) CG: M = 52.7 (24.7) <b>p &lt; 0.01</b> Cohens d = 0.47 [0.13–0.81]</p> <p><b>t2 (6 months follow-up)</b></p> <p>IG: M = 41.9 (30.0)</p>
	reduction of therapy-related efforts and strains for patients/their relatives	general psychopathology	Brief Symptom Inventory (BSI)	Short version of the Symptom Checklist-90 (SCL-90)		<p>IG: M = 1.34 (0.56) CG: M = 1.27 (0.57)</p> <p>IG: M = 0.94 (0.63) CG: M = 1.18 (0.71) <b>p &lt; 0.001</b> Cohens d = 0.42 [0.08–0.75]</p> <p>IG: M = 0.97 (0.77)</p>
		anxiety-related symptoms	Beck Anxiety Inventory (BAI)	0–63 higher scores indicate stronger symptoms 0–21 = low anxiety 22–35 = moderate anxiety > 36 = potentially concerning levels of anxiety [46]		<p>IG: M = 34.9 (9.1) CG: M = 33.3 (10.3)</p> <p>IG: M = 27.8 (9.1) CG: M = 31.4 (10.0) <b>p &lt; 0.001</b> Cohens d = 0.41 [0.07–0.74]</p> <p>IG: 26.6 (9.4)</p>

Table 5 (continued)

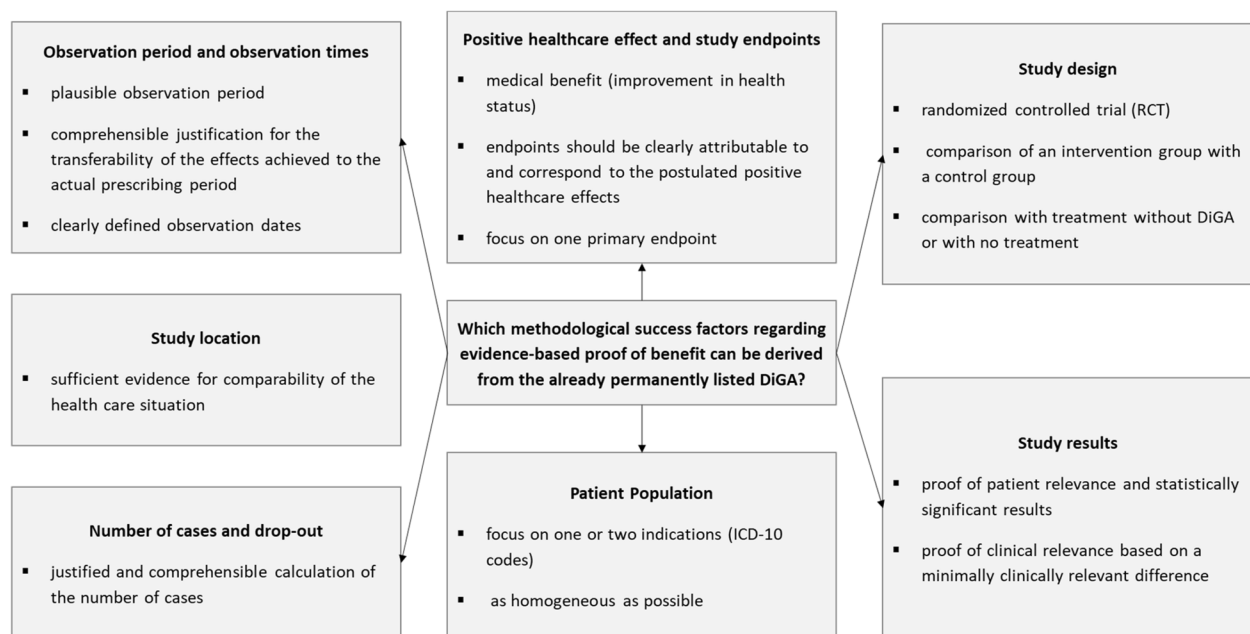
DiGA	Positive care effect	Primary endpoint	Operationalization	Interpretation	MCID	Results
		depressive symptoms	Beck Depression Inventory-II (BDI-II)	0–63 higher scores indicate stronger symptoms 0–8 = no depression 9–13 = minimal depression 14–19 = mild depression 20–28 = moderate depression 29–63 = severe depression [46, 47]		IG: M = 22.6 (10.6) CG: M = 22.0 (11.0) <b>p &lt; 0.001</b> Cohens d = 0.61 [0.27–0.95] IG: M = 15.8 (12.4) CG: M = 22.9 (12.6) IG: M = 16.3 (13.7)
		quality of life	SF-12 Mental Health	0–100 higher scores indicate higher quality of life [52]		IG: M = 31.2 (8.8) CG: M = 33.2 (9.5) <b>p &lt; 0.001</b> Cohens d = 0.49 [0.15–0.83] IG: M = 37.5 (11.8) CG: M = 33.0 (9.2) IG: M = 39.9 (12.2)
			SF-12 Physical Health	0–100 higher scores indicate higher quality of life [52]		IG: M = 48.5 (11.2) CG: M = 48.3 (10.8) <b>p = 0.19</b> Cohens d = 0.16 [–0.17–0.50] IG: M = 48.3 (11.4) CG: M = 47.2 (9.5) IG: M = 48.6 (11.1)
<b>Vivira [32]</b>	improvement of the state of health	back, knee and hip pain	visual numeric rating scale			<b>t0 (Baseline)</b> <b>t1 (2 weeks)</b> <b>t2 (6 weeks)</b> <b>t3 (12 weeks)<sup>a</sup></b> difference in change from baseline = -2.34 [-2.84; -1.83] <b>p &lt; 0.0001</b> difference in change from baseline = -2.44 [-2.92; -1.95] <b>p &lt; 0.0001</b>
<b>Vorvida [33]</b>	improvement of the state of health	daily average consumption of pure alcohol in grams)	Quantity-frequency index: self-reported quantities of alcohol in g (last 30 days)			<b>t0 (Baseline)</b> <b>t1 (12 weeks)<sup>a</sup></b> <b>t2 (6 months)</b> IG: M = 63.69 (61.84) CG: M = 61.64 (58.84) IG: mM = 40.8 (3.3) CG: mM = 56.8 (3.3) <b>p = 0.001</b> Cohens d = 0.278 <b>p &lt; 0.001</b> Cohens d = 0.327

**Table 5** (continued)

DIGA	Positive care effect	Primary endpoint	Operationalization	Interpretation	MCID	Results
			Timeline-Follow-Back: self-reported amount of alcohol in g (last 7 days)			IG: mM = 52.91 (56.68) CG: M = 46.82 (41.18)  IG: mM = 34.3 (1.3) CG: mM = 43.7 (1.3) <b>p &lt; 0.001</b> Cohens d = 0.419  <b>t0 (Baseline)</b>  <b>t1 (9 months)<sup>a</sup></b> MD = -5.67% <b>p = 0.001</b> 95% CI [-7.32, -4.02]  <b>t2 (12 months)</b> MD = -7.75% <b>p = 0.001</b> 95% CI [-9.61, -5.89]
Zanadio [34]	improvement of the state of health	weight reduction of at least 5%	BMI			

Unless otherwise stated, all values are given as: mean/marginal mean/mean difference (standard deviation), effect size [95% confidence interval]  
CI Confidence interval, DIGA Digital health application, MCID Minimal clinically important difference, IG Intervention group, CG Control group, M Mean, mM Marginal mean, MD Mean difference, ITT Intention-to-treat, PP Per protocol, n.r. Not reported

<sup>a</sup> Survey time of the primary endpoint



**Fig. 2** Overview of methodological success factors regarding evidence-based proof of benefit of permanently listed DiGA

DiGA = Digital health application, ICD = International Classification of Diseases

transparency, which is relevant to further development of the DiGA approval process.

First, of the four DiGA initially demonstrating positive healthcare effects through systematic data evaluation, hardly any pilot study results were available. Only the Selfapy applications provisional results could be found in the DiGA directory.

Second, the DiGA directory was not sufficient for a complete analysis of the seven identified categories and the evidence produced; further sources, such as study registries, study reports and publications, had to be obtained. Finally, we found one case in which the results reported in the DiGA directory did not match to those from the study report.

## Discussion

The aim of this study was to identify BfArMs requirements of studies that prove a postulated positive healthcare effect, and, to assess the evidence of applications permanently listed in the DiGA directory.

Permanently listed DiGA meet higher requirements for demonstrating a positive healthcare effect, than those specified in the guideline. Most DiGA focused on one or two indications (ICD-10 codes), and proved a medical benefit through a RCT. All DiGA studies were able to prove patient relevance and achieved statistically significant results, however, observation periods, sample sizes and drop-out rates differed substantially among studies.

## Patient population and patient characteristics

Permanently listed mental health DiGA are most frequently represented in the DiGA directory. A focus on one or two indications (ICD-10 codes) seems to predict success for a permanent listing. This could be explained by specific and homogeneous patient populations being needed to provide adequate evidence. Surprising is however, that the BfArM guideline clearly specifies that inclusion in the DiGA directory is only possible for three or four-digit ICD-10 codes. Nevertheless, three manufacturers (Hello Better Panik, Velibra, Zanadio) achieved permanent listing using only a five-digit ICD-10 code.

In all studies aiming to demonstrate a positive healthcare effect, the proportion of women represented in the populations (in both IG and CG) was higher than that of men. A report from a German health insurance company also showed that more women use DiGA than men. Since gender-specific disease prevalence among insured patients in the (as of yet) main DiGA indications can sometimes differ greatly, this may be a result [53].

In almost all studies, the mean age of the participants was between 35 and 55 years. An analysis by a German health insurer of the age structure of all insured adults with at least one DiGA claim as of December 31<sup>st</sup>, 2021 showed that 27% were aged 50–59 years, followed by 22% aged 30–39 years, and 20% aged 40–49 [53].

DiGA studies published to date show that individuals with better education and employment status participated

more frequently. International studies also show that the use of mobile apps is unevenly distributed in society; a US-America study showed that people with an academic education had a higher (2.8-fold) chance of using a health app, compared to those without a high school diploma. The study also showed that higher education was associated with more frequent use of digital health services [54]. A study from the Netherlands also points to social differences in the use of health apps, and showed that people who used a health app were younger and more educated, compared to those who did not [55]. An important factor that should be considered in relation to these discussed aspects is the willingness to participate in a clinical trial. Gouveia et al. (2022) showed that the willingness of younger patients to participate was significantly higher compared to older patients. In contrast, gender, lifestyle, employment status, monthly income, or education showed no influence on willingness to participate [56]. Another point influencing participation in a clinical trial is personal treatment preference [57]. In addition to willingness, attitude, and patients' motivation to use a DiGA and/or participate in a trial, their previous experience with digital tools and their habits to use such tools for their disease self-management needs to be considered [58, 59]. Due to the influencing factors mentioned above, and the scarce data on DiGA implementation, it is not fully understood whether DiGA address the entire defined study population, or only certain subgroups. This fact underlines the high relevance of subgroup analyses for the manufacturers.

There is a need for further research on the monitoring of the actual use of DiGA by all statutory health insured in Germany, to assess whether the study populations correspond to the actual target groups. At the same time, during the process of the DiGA approval, responsible authorities should monitor whether pilot/feasibility studies, and those needed for permanent listing, may have the potential to recruit more diverse populations, and thereby increase external validity of the trials. The inherent concern is that digitally competent and literate patient populations are recruited, and other specific populations overlooked. This may further amplify the digital divide in Germany [60]. Also Stern et al. (2022) point out that equity aspects must be taken into account to ensure that health inequalities are not reinforced, or even created in the first place [12]. As such, DiGA manufacturers should be encouraged to take additional efforts to effectively recruit all genders, as well as a variety of cultural backgrounds and different levels of digital health literacy. This would enable subgroup analyses, identify those requiring additional support for adequate DiGA usage, and ultimately support individualized DiGA recommendations, according to patient characteristics, competencies and preferences [61].

### Positive healthcare effect and study endpoints

It is striking that— although patient-relevant structural and procedural improvements show high potential for improving care, especially in terms of processes — all DiGA have provided a positive care effect via a medical benefit. Lantzsch et al. (2022) pointed out needed improvement in using outcomes for patient-relevant improvement of structure and processes regarding an improved evaluation [11]. All applications were able to demonstrate an improvement in health status. Only one DiGA additionally aimed for an improved quality of life, and a reduced disease duration, which, however, could not be achieved. DiGA that provide a medical benefit in the form of a health status improvement seem to have the greatest likelihood to be permanently listed in the DiGA directory. Focusing on one primary endpoint seems to predict success for a permanent listing; the majority of the DiGA provide evidence for one defined primary endpoint. With regards to the study endpoints and measurements used, the analysis found that multiple endpoints were used, which is not surprising, given the range of indications covered. However, the heterogeneous measurements used to assess these endpoints offer a standardization potential, which will be increasingly relevant as more DiGA for the same indications are listed. To close this gap, it is possible to use Core Outcome Sets (COS). The idea of COS, is to provide a minimum set of outcome domains and a consensual set of measurement tools to be used in every clinical study with a comparable intervention and target population. Development of new, or expansion of existing COS for the evaluation of DiGA, can improve comparability of study results [62, 63].

### Study design

Although BfArM accepts study designs with a lower level of evidence for the proof of a positive healthcare effect, all manufacturers of permanently listed DiGA conducted a study with a high level of evidence. Implementing a RCT was therefore a factor increasing the likelihood of permanent inclusion in the DiGA directory. RCTs are not only conducted in the medical sector, but increasingly also in the technology sector [64], and are therefore considered a promising study design for DiGA. However, the comparatively short innovation cycles for new technologies can be an obstacle to conducting RCTs in a DiGA context. The DiGA and its individual components are usually continuously adapted and further developed by the manufacturers, so that new versions are often already available before the evaluation of the original version is completed. One success factor here can be a continuous, learning evaluation of the continuously changing DiGA process [14]. Stern et al. (2022) cite actual or perceived risks of regulatory uncertainty, as possible reasons for

deciding against non-RCT studies. It may then be more risky or costly for manufacturers to subsequently switch to a more traditional study design after an unsuccessful trial [12].

Although both inter- and intra-individual comparisons are allowed, all manufacturers chose to compare an intervention group with a control group. Different methods are proposed for this comparison. DiGA manufacturers have compared to treatment without DiGA, or to no treatment. The third option — comparison with another comparable DiGA that is already permanently listed in the DiGA directory at the time of application — will become more relevant as the DiGA directory grows. Increasing in popularity is a fourth conceivable pathway; the design of a placebo DiGA, allowing blinding to be maintained. Stern et al. (2022) emphasize the relevance of establishing best practices and methods to accurately define the comparison group [12].

#### Observation period and observation times

BfArM requires a description of baseline data, data collection periods, and follow-up times, but leave manufacturers flexibility in observation periods. Despite this, choosing a plausible observation period fitting to the indication, is important. If the observation period is shortened, justifying transferability of the achieved effect(s) to the prescribing period, is advisable. Stern et al. (2022) consider planning of washout phases, in which patients do not receive any therapies prior to the start of the actual intervention [12]. In any case, the observation dates must be clearly defined. In fact, almost all studies date back several years, meaning already existing data were used to prove the positive health effect. Since May 27<sup>th</sup>, 2020, manufacturers can submit applications to be included in the DiGA directory. The directory was started on October 6<sup>th</sup>, 2020. Only three trials (Kalmeda, Vivira, Zanadio) started in 2020/2021. The study for the application Vivira led to permanent listing in the DiGA directory in only 12 weeks, with a total study duration of five months.

#### Study results

According to BfArM, a positive healthcare effect is considered proven if the outcome is (clinically) relevant, patient-relevant, and statistically significant. All DiGA study results were able to prove patient relevance and achieve statistical significance, and are central to successful entry in the registry.

Although the questionnaire scores improved after the application of the DiGA, the symptoms partly remained similar. One example is Hello Better Panik; here, though participants' scores on the PAS scale (German version: Panik- und Agoraphobieskala) improved by 6.45 (IG)

and 1.83 (CG) from baseline (compared to the primary endpoint), symptoms remained moderate. Similarly, Hello Better Stress und Burnout showed improved Perceived Stress Scale scores (by 8.01 (IG) and 2.19 (CG) from baseline compared to the primary endpoint), however symptoms also remained moderate. These examples show the limitations of DiGA. Somnio, however, showed differences in the Insomnia Severity Index of 7.58 (IG) and 1.22 (CG), thus improving from moderate insomnia to subthreshold insomnia; DiGA also have the potential to greatly improve users' health.

#### Further results and recommendations

In comparing the evidence required with the evidence provided by permanently listed DiGA, it was clear that transparency is not yet fully apparent. Particularly, gaining a fully comprehensive picture on some aspects (patient population, study endpoints, study location, drop-out rates, and study results) was not possible, using the DiGA directory alone. Other sources had to be consulted, such as study registers, study reports and publications. In one case, the results reported in the DiGA directory did not match those in the study report. The claimed user-friendliness of the directory must also be questioned. The user only gets an insight into the mean values/mean value comparisons. An interpretation aid for the meaning of these figures is missing.

In the future, the DiGA directory should be an essential tool for medical professionals. An important step for the spread of health apps and their integration into clinical practice, is the education of clinicians regarding available technologies [4]. A study with 51 physicians showed that half of the respondents expect to be able to identify high-quality DiGA using the DiGA directory. This group is more likely to prescribe DiGA [65]. To increase DiGA acceptance and willingness to prescribe among physicians, a transparent, complete, and correct presentation of the information is indispensable.

Future research needs are identified in the continuous evaluation of the fast-track process, with special attention to requirements of DiGA manufacturers. As an innovative and learning system in healthcare which is constantly being developed and improved (with increasingly specific requirements for DiGA manufacturers, and the growing DiGA directory), an accompanying monitoring could be a central success factor.

#### Limitations

The current study combined multiple methods to analyze the available evidence of listed DiGA. However, only available studies, data, and information on the applications permanently listed in the DiGA directory were included in the analysis, due to the lack of



evidence for provisionally listed applications. Therefore, recommendations may not be transferable to DiGA manufacturers aiming for provisional listing. Furthermore, evidence available for permanently listed DiGA was not identified through a systematic literature search. To limit the risk of overlooking relevant studies, numerous independent sources were used to identify DiGA-specific information. In addition to the DiGA directory, entries in study registries, study protocols, study reports, submitted publications, and manufacturers' websites, were used as information sources.

## Conclusion

The results of this analysis indicate that permanently listed DiGA meet higher standards than required by the guideline. Before prescribing a DiGA, its evidence should be carefully examined. The identified success factors provide healthcare practitioners with a transparent overview of the status quo of applications already tested. They can also support future manufacturers in the development and evaluation of their DiGA. With regards to the various endpoints used in the presented studies, future DiGA trials should focus on the most relevant outcomes, and strive towards comparability of results (especially among DiGA for the same indication). In addition, an analysis of DiGA-use in everyday care should become the subject of further implementation research. Overall, there is a need for accompanying monitoring; from the application development, through its testing during the evaluation study, to its use in everyday care. A growing DiGA system could be a beneficial improvement for the whole healthcare system.

## Abbreviations

BfArM	Bundesinstitut für Arzneimittel- und Medizinprodukte
CG	Control group
COS	Core Outcome Set
DiGA	Digital health application
DRKS	Deutsches Register klinischer Studien
ICD-10	International Classification of Diseases
IG	Intervention group
MDR	Medical Device Regulation
RCT	Randomized controlled trial

## Acknowledgements

We thank Joanna Dising for the linguistic optimization of our manuscript.

## Authors' contributions

MM created the concept and the methodology of the manuscript, wrote the main manuscript text, and created the figures and tables. DH, PT and TS assisted in the conceptualization and preparation of the manuscript. CMH, SS and RH assisted in quality assurance. All authors reviewed the manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL. Funded by the Open Access Publishing Fund of Leipzig University, which is supported by the German Research Foundation within the program Open Access Publication Funding.

## Availability of data and materials

All data generated or analyzed during this study are included in this published article:

DiGA	Direct web links
Deprexis	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/450">https://diga.bfarm.de/de/verzeichnis/450</a> Publications: Meyer et al. (2015): <a href="https://www.jmir.org/2009/2/e15/PDF">https://www.jmir.org/2009/2/e15/PDF</a> Klein et al. (2016): <a href="https://www.karger.com/Article/Pdf/445355">https://www.karger.com/Article/Pdf/445355</a>
Elevida	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/419">https://diga.bfarm.de/de/verzeichnis/419</a> Publication: Pöttgen et al. (2018): <a href="https://www.researchgate.net/publication/323820297_Randomised_controlled_trial_of_a_self-guided_online_fatigue_intervention_in_multiple_sclerosis">https://www.researchgate.net/publication/323820297_Randomised_controlled_trial_of_a_self-guided_online_fatigue_intervention_in_multiple_sclerosis</a> Study report: Mayer et al. (2020): <a href="https://elevida.de/downloads/studienbericht_elevida_poettgen_2018.pdf">https://elevida.de/downloads/studienbericht_elevida_poettgen_2018.pdf</a>
Hello Better Diabetes and Depression	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/1376">https://diga.bfarm.de/de/verzeichnis/1376</a> Publication: Nobis et al. (2015): <a href="https://diabetesjournals.org/care/article/38/5/776/37466/Efficacy-of-a-Web-Based-Intervention-With-Mobile">https://diabetesjournals.org/care/article/38/5/776/37466/Efficacy-of-a-Web-Based-Intervention-With-Mobile</a> Study report: Balzus et al. (2021): <a href="https://helloworld.de/wp-content/uploads/2021/08/Studienbericht_Diabetes.pdf">https://helloworld.de/wp-content/uploads/2021/08/Studienbericht_Diabetes.pdf</a> Study protocol: Nobis et al. (2013): <a href="https://bmcp psychiatry.biomedcentral.com/articles/10.1186/1471-244X-13-306">https://bmcp psychiatry.biomedcentral.com/articles/10.1186/1471-244X-13-306</a>
Hello Better Panik	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/1513">https://diga.bfarm.de/de/verzeichnis/1513</a> Publication: Ebenfeld et al. (2021): <a href="https://www.jmir.org/2021/3/e20829/PDF">https://www.jmir.org/2021/3/e20829/PDF</a> Study report: Feiler et al. (2021): <a href="https://docplayer.org/228425224-Bezeichnung-der-klinischen-studie-smart-phonebasiertes-online-training-zur-bewaeltigung-von-panikatacken-und-agoraphobie.html">https://docplayer.org/228425224-Bezeichnung-der-klinischen-studie-smart-phonebasiertes-online-training-zur-bewaeltigung-von-panikatacken-und-agoraphobie.html</a> Study protocol: Ebenfeld et al. (2014): <a href="https://trialsjournal.biomedcentral.com/articles/10.1186/1745-6215-15-427">https://trialsjournal.biomedcentral.com/articles/10.1186/1745-6215-15-427</a>
Hello Better Stress and Burnout	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/965">https://diga.bfarm.de/de/verzeichnis/965</a> Publication: Heber et al. (2016): <a href="https://www.jmir.org/2016/1/e21/PDF">https://www.jmir.org/2016/1/e21/PDF</a> Study report: Feiler et al. (2021): <a href="https://helloworld.de/wp-content/uploads/2021/06/Studienbericht_Hello_Better_Stress_und_Burnout.pdf">https://helloworld.de/wp-content/uploads/2021/06/Studienbericht_Hello_Better_Stress_und_Burnout.pdf</a> Study protocol: Heber et al. (2013): <a href="https://bmcp public health.biomedcentral.com/articles/10.1186/1471-2458-13-655">https://bmcp public health.biomedcentral.com/articles/10.1186/1471-2458-13-655</a>



DiGA	Direct web links
Hello Better Vaginismus Plus	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/1497">https://diga.bfarm.de/de/verzeichnis/1497</a> Publication: Zarski et al. (2021): <a href="https://www.researchgate.net/publication/356625888_Efficacy_of_internet-based_treatment_for_genito-pelvic_painpenetration_disorder_Results_of_a_randomized_controlled_trial">https://www.researchgate.net/publication/356625888_Efficacy_of_internet-based_treatment_for_genito-pelvic_painpenetration_disorder_Results_of_a_randomized_controlled_trial</a> Study report: Feiler et al. (2021): <a href="https://hellobetter.de/wp-content/uploads/2021/10/Studienbericht-Vaginismus-Plus.pdf">https://hellobetter.de/wp-content/uploads/2021/10/Studienbericht-Vaginismus-Plus.pdf</a> Study protocol: Zarski et al. (2018): <a href="https://www.frontiersin.org/articles/10.3389/fpsy.2017.00260/full">https://www.frontiersin.org/articles/10.3389/fpsy.2017.00260/full</a>
Kalmeda	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/350">https://diga.bfarm.de/de/verzeichnis/350</a> Study report: Stover et al. (2022): <a href="https://drks.de/search/de/trial/DRKS00022973">https://drks.de/search/de/trial/DRKS00022973</a>
Selfapy Depression	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/876">https://diga.bfarm.de/de/verzeichnis/876</a> Publication: Krämer et al. (2022): <a href="https://formative.jmir.org/2022/4/e34330/PDF">https://formative.jmir.org/2022/4/e34330/PDF</a>
Somnio	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/508">https://diga.bfarm.de/de/verzeichnis/508</a> Publication: Lorenz et al. (2019): <a href="https://www.zora.uzh.ch/id/eprint/157067/1/randomized_controlled_trial_to_test_the_efficacy_of_an_unguided_online_intervention_with_automated_feedback_for_the_treatment_of_insomnia.pdf">https://www.zora.uzh.ch/id/eprint/157067/1/randomized_controlled_trial_to_test_the_efficacy_of_an_unguided_online_intervention_with_automated_feedback_for_the_treatment_of_insomnia.pdf</a>
Velibra	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/316">https://diga.bfarm.de/de/verzeichnis/316</a> Publication: Berger et al. (2016): <a href="https://boris.unibe.ch/94877/1/effects_of_a_transdiagnostic_unguided_internet_intervention_velibra_for_anxiety_disorders_in_primary_care_results_of_a_randomized_controlled_trial.pdf">https://boris.unibe.ch/94877/1/effects_of_a_transdiagnostic_unguided_internet_intervention_velibra_for_anxiety_disorders_in_primary_care_results_of_a_randomized_controlled_trial.pdf</a> Study report: Mayer et al. (2020): <a href="https://de.velibra.com/downloads/Studienbericht_velibra_Berger_2017_inkl_Append_2020-06-26.pdf">https://de.velibra.com/downloads/Studienbericht_velibra_Berger_2017_inkl_Append_2020-06-26.pdf</a>
Vivira	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/387">https://diga.bfarm.de/de/verzeichnis/387</a>
Vorvida	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/868">https://diga.bfarm.de/de/verzeichnis/868</a> Publication: Zill et al. (2019): <a href="https://www.aerzteblatt.de/archiv/205622/Wirksamkeit-einer-Internetintervention-zur-Reduktion-von-Alkoholkonsum-bei-Erwachsenen">https://www.aerzteblatt.de/archiv/205622/Wirksamkeit-einer-Internetintervention-zur-Reduktion-von-Alkoholkonsum-bei-Erwachsenen</a> Study report: Mayer et al. (2020): <a href="https://de.vorvida.com/downloads/Studienbericht_Zill_2019_vorvida_20201105.pdf">https://de.vorvida.com/downloads/Studienbericht_Zill_2019_vorvida_20201105.pdf</a> Study protocol: Zill et al. (2016): <a href="https://bmcp psychiatry.biomedcentral.com/articles/10.1186/s12888-016-0725-9">https://bmcp psychiatry.biomedcentral.com/articles/10.1186/s12888-016-0725-9</a>
Zanadio	DiGA directory: <a href="https://diga.bfarm.de/de/verzeichnis/294">https://diga.bfarm.de/de/verzeichnis/294</a>

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>Faculty of Economics and Management Science, Leipzig University, Chair for Health Economics and Management, Leipzig, Germany. <sup>2</sup>Scientific Institute for Health Economics and Health System Research (WIG2 GmbH), Markt 8, 04109 Leipzig, Germany. <sup>3</sup>Department of Health Sciences/Public Health, Dresden University, Dresden, Germany.

Received: 5 December 2022 Accepted: 15 March 2023

Published online: 17 April 2023

## References

- Essén A, Stern AD, Haase CB, et al. Health app policy: international comparison of nine countries' approaches. *NPJ Digit Med*. 2022;5(1):31. <https://doi.org/10.1038/s41746-022-00573-1>.
- Giebel GD, Speckemeier C, Abels C, et al. Problems and Barriers Related to the Use of Digital Health Applications: Protocol for a Scoping Review. *JMIR Res Protoc*. 2022;11(4):e32702. <https://doi.org/10.2196/32702>.
- BfArM. Das Fast-Track-Verfahren für digitale Gesundheitsanwendungen (DiGA) nach § 139e SGB V: Ein Leitfaden für Hersteller, Leistungserbringer und Anwender. 2022. Available at [https://www.bfarm.de/SharedDocs/Downloads/DE/Medizinprodukte/diga\\_leitfaden.html](https://www.bfarm.de/SharedDocs/Downloads/DE/Medizinprodukte/diga_leitfaden.html).
- Gordon WJ, Landman A, Zhang H, et al. Beyond validation: getting health apps into clinical practice. *NPJ Digit Med*. 2020;3:14. <https://doi.org/10.1038/s41746-019-0212-z>.
- WHO. Global strategy on digital health 2020–2025 2021. Available at <https://apps.who.int/iris/bitstream/handle/10665/344249/978924002924-eng.pdf>. Accessed 11 July, 2022.
- Ryll B. Digitale Gesundheitsanwendungen (DiGA): Patientenzentrierte Gesundheitsversorgung mit disruptivem Potenzial. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitschutz*. 2021;64(10):1207–12. <https://doi.org/10.1007/s00103-021-03421-x>.
- Brönneke JB, Hagen J, Kircher P, et al. Digitized healthcare in 2030-a possible scenario. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitschutz*. 2021;64(10):1285–91. <https://doi.org/10.1007/s00103-021-03416-8>.
- Heimann P, Lorenz N, Blum N, et al. Experiences of digital health care applications (DiGA) manufacturers with the BfArM Fast-Track procedure. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitschutz*. 2021;64(10):1249–53. <https://doi.org/10.1007/s00103-021-03422-w>.
- Löbker W, Böhmer AC, Höfgen B. Support for innovation at the BfArM-experiences from the consultations on digital health applications (DiGA). *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitschutz*. 2021;64(10):1241–8. <https://doi.org/10.1007/s00103-021-03410-0>.
- Düvel JA, Gensorowsky D, Hasemann L, et al. Lösungsansätze für den Zugang digitaler Gesundheitsanwendungen zur Gesetzlichen Krankenversicherung: eine qualitative Studie. *Gesundheitswesen (Bundesverband der Ärzte des Öffentlichen Gesundheitsdienstes (Germany))* 2021. Available at <https://pubmed.ncbi.nlm.nih.gov/33636736/>.
- Lantzsch H, Eckhardt H, Campione A, et al. Digital health applications and the fast-track pathway to public health coverage in Germany: Challenges and opportunities based on first results. *BMC Health Serv Res*. 2022;22(1):1–6.
- Stern AD, Brönneke J, Debatin JF, et al. Advancing digital health applications: priorities for innovation in real-world evidence generation. *Lancet Digit Health*. 2022;4(3):e200–6. [https://doi.org/10.1016/S2589-7500\(21\)00292-2](https://doi.org/10.1016/S2589-7500(21)00292-2).

13. Geier AS. Digitale Gesundheitsanwendungen (DiGA) auf dem Weg zum Erfolg – die Perspektive des Spitzenverbandes Digitale Gesundheitsversorgung. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitschutz*. 2021;64(10):1228–31. <https://doi.org/10.1007/s00103-021-03419-5>.
14. Hemkens LG. Nutzenbewertung digitaler Gesundheitsanwendungen – Herausforderungen und Möglichkeiten. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitschutz*. 2021;64(10):1269–77. <https://doi.org/10.1007/s00103-021-03413-x>.
15. Jandoo T. WHO guidance for digital health: What it means for researchers. *Digit Health*. 2020;6:2055207619898984. <https://doi.org/10.1177/2055207619898984>.
16. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.3 (updated February 2022). Cochrane. 2022. Available at [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook).
17. Meyer B, Berger T, Caspar F, et al. Effectiveness of a novel integrative online treatment for depression (Deprexis): randomized controlled trial. *J Med Internet Res*. 2009;11(2):e15. <https://doi.org/10.2196/jmir.1151>.
18. Klein JP, Berger T, Schröder J, et al. Effects of a Psychological Internet Intervention in the Treatment of Mild to Moderate Depressive Symptoms: Results of the EVIDENT Study, a Randomized Controlled Trial. *Psychother Psychosom*. 2016;85(4):218–28.
19. Pöttgen J, Moss-Morris R, Wendebourg J-M, et al. Randomised controlled trial of a self-guided online fatigue intervention in multiple sclerosis. *J Neurol Neurosurg Psychiatry*. 2018;89(9):970–6.
20. Nobis S, Lehr D, Ebert DD, et al. Efficacy of a web-based intervention with mobile phone support in treating depressive symptoms in adults with type 1 and type 2 diabetes: a randomized controlled trial. *Diabetes Care*. 2015;38(5):776–83.
21. Balzus L, Feiler M, Dr. Stephani V, et al. Studienbericht: HelloBetter Diabetes und Depression. 2021.
22. Ebenfeld L, Kleine Stegemann S, Lehr D, et al. Efficacy of a hybrid online training for panic symptoms and agoraphobia: study protocol for a randomized controlled trial. *Trials*. 2021;15:427.
23. Feiler M, Balzus L, Groos H, et al. Studienbericht: Hello Better Panik. 2021.
24. Heber E, Lehr D, Ebert DD, et al. Web-Based and Mobile Stress Management Intervention for Employees: A Randomized Controlled Trial. *J Med Internet Res*. 2016;18(1):e21.
25. Feiler M, Balzus L, Prof. Dr. DD, Ebert, et al. Studienbericht: Hello Better Stress und Burnout. 2021.
26. Zarski A-C, Berking M, Ebert DD. Efficacy of internet-based treatment for genito-pelvic pain/penetration disorder: Results of a randomized controlled trial. *J Consult Clin Psychol*. 2021;89(11):909–24.
27. Feiler M, Balzus L, Groos H, et al. Studienbericht: Hello Better Vaginismus Plus. 2021.
28. Stover C. Kalmeda Tinnitus-Studie: Abschlussbericht 2022. Available at [http://www.kalmeda.de/sites/default/files/files/2022-10/Abschlussbericht\\_Kalmeda\\_V1.0-2022-09-26\\_0.pdf](http://www.kalmeda.de/sites/default/files/files/2022-10/Abschlussbericht_Kalmeda_V1.0-2022-09-26_0.pdf). Accessed 23 Oct 2022.
29. Krämer R, Köhler S. Evaluation of the online-based self-help programme "Selfapy" in patients with unipolar depression: study protocol for a randomized, blinded parallel group dismantling study. *Trials*. 2022;22(1):264.
30. Lorenz N, Heim E, Roetger A, et al. Randomized Controlled Trial to Test the Efficacy of an Unguided Online Intervention with Automated Feedback for the Treatment of Insomnia. *Behav Cogn Psychother*. 2019;47(3):287–302.
31. Berger T, Urech A, Krieger T, et al. Effects of a transdiagnostic unguided Internet intervention ('velibra') for anxiety disorders in primary care: results of a randomized controlled trial. *Psychol Med*. 2017;47(1):67–80.
32. BfArM. DiGA-Verzeichnis: Vivira. Available at <https://diga.bfarm.de/de/verzeichnis/387>. Accessed 24 Oct 2022.
33. Zill JM, Christalle E, Meyer B, et al. The Effectiveness of an Internet Intervention Aimed at Reducing Alcohol Consumption in Adults. *Dtsch Arztebl Int*. 2019;116(8):127–33.
34. BfArM. DiGA-Verzeichnis: Zanadio. Available at <https://diga.bfarm.de/de/verzeichnis/294>. Accessed 24 Oct 2022.
35. Löwe B, Unützer J, Callahan CM, et al. Monitoring depression treatment outcomes with the patient health questionnaire-9. *Med Care*. 2004;42(12):1194–201. <https://doi.org/10.1097/00005650-200412000-00006>.
36. Mayer J, Meyer B, Bültmann O. Klinischer Studienbericht "elevida." 2020.
37. Pouchot J, Kherani RB, Brant R, et al. Determination of the minimal clinically important difference for seven fatigue measures in rheumatoid arthritis. *J Clin Epidemiol*. 2008;61(7):705–13. <https://doi.org/10.1016/j.jclinepi.2007.08.016>.
38. Goligher EC, Pouchot J, Brant R, et al. Minimal clinically important difference for 7 measures of fatigue in patients with systemic lupus erythematosus. *J Rheumatol*. 2008;35(4):635–42.
39. Radloff LS. The CES-D Scale. *Appl Psychol Meas*. 1977;1(3):385–401. <https://doi.org/10.1177/014662167700100306>.
40. Hautzinger M, Bailer M, Hofmeister D, et al. *Allgemeine Depressionsskala*. 2. überarbeitete, neu normierte Auflage. Göttingen: Hogrefe; 2012. Available at <https://www.testzentrale.de/shop/allgemeine-depressionsskala.html>.
41. Cuijpers P, Turner EH, Koole SL, et al. What is the threshold for a clinically relevant effect? The case of major depressive disorders. *Depress Anxiety*. 2014;31(5):374–8. <https://doi.org/10.1002/da.22249>.
42. Bandelow B. Assessing the efficacy of treatments for panic disorder and agoraphobia. II. The Panic and Agoraphobia Scale. *Int Clin Psychopharmacol*. 1995;10(2):73–81. <https://doi.org/10.1097/00004850-19950600000003>.
43. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav*. 1983;24(4):385–96.
44. van Lankveld JJDM, ter Kuile MM, de Groot HE, et al. Cognitive-behavioral therapy for women with lifelong vaginismus: a randomized waiting-list controlled trial of efficacy. *J Consult Clin Psychol*. 2006;74(1):168–78. <https://doi.org/10.1037/0022-006X.74.1.168>.
45. Hiller W, Goebel G. Rapid assessment of tinnitus-related psychological distress using the Mini-TQ. *Int J Audiol*. 2004;43(10):600–4. <https://doi.org/10.1080/14992020400050077>.
46. Beck AT, Epstein N, Brown G, et al. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol*. 1988;56(6):893–7. <https://doi.org/10.1037//0022-006X.56.6.893>.
47. Beck AT, Steer RA, Ball R, et al. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *J Pers Assess*. 1996;67(3):588–97. [https://doi.org/10.1207/s15327752jpa6703\\_13](https://doi.org/10.1207/s15327752jpa6703_13).
48. Bastien C. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med*. 2001;2(4):297–307 Available at <https://pubmed.ncbi.nlm.nih.gov/11438246/>.
49. Mayer J, Meyer B, Bültmann O. Klinischer Studienbericht "velibra". 2020.
50. Lovibond SH, Lovibond PF. *Manual for the depression anxiety stress scales*. 2nd ed. Sydney: Psychology Foundation of Australia; 1996.
51. Ronk FR, Korman JR, Hooke GR, et al. Assessing clinical significance of treatment outcomes using the DASS-21. *Psychol Assess*. 2013;25(4):1103–10. <https://doi.org/10.1037/a0033100>.
52. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34(3):220–33. <https://doi.org/10.1097/00005650-199603000-00003>.
53. Techniker Krankenkasse. DiGA-Report 2022. 2022.
54. Bhuyan SS, Lu N, Chandak A, et al. Use of Mobile Health Applications for Health-Seeking Behavior Among US Adults. *J Med Syst*. 2016;40(6):153. <https://doi.org/10.1007/s10916-016-0492-7>.
55. Bol N, Helberger N, Weert JCM. Differences in mobile health app use: A source of new digital inequalities? *Inf Soc*. 2018;34(3):183–93. <https://doi.org/10.1080/01972243.2018.1438550>.
56. Gouveia R, Cruz VT and Almeida L. Sociodemographic and psychological characteristics influencing patients' willingness to participate in clinical trials. *BMJ Open Qual* 2022;11(4). <https://doi.org/10.1136/bmjopen-2022-002044>.
57. Sheridan R, Martin-Kerry J, Hudson J, et al. Why do patients take part in research? An overview of systematic reviews of psychosocial barriers and facilitators. *Trials*. 2020;21(1):259. <https://doi.org/10.1186/s13063-020-4197-3>.
58. Uncovska M, Freitag B, Meister S, et al. Patient Acceptance of Prescribed and Fully Reimbursed mHealth Apps in Germany: An UTAUT2-based Online Survey Study. *J Med Syst*. 2023;47(1):14. <https://doi.org/10.1007/s10916-023-01910-x>.
59. Venkatesh V, Thong JY, Xu X. Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Q*. 2012;36(1):157. <https://doi.org/10.2307/41410412>.

60. Müller CA, Wachtler B, Lampert T. Digital Divide – Soziale Unterschiede in der Nutzung digitaler Gesundheitsangebote. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2020;63(2):185–91. <https://doi.org/10.1007/s00103-019-03081-y>.
61. Rivera-Romero O, Gabarron E, Miron-Shatz T, et al. Social Media, Digital Health Literacy, and Digital Ethics in the Light of Health Equity. *Yearb Med Inform*. 2022. <https://doi.org/10.1055/s-0042-1742503>.
62. Williamson PR, Altman DG, Bagley H, et al. The COMET Handbook: version 1.0. *Trials*. 2017;18(Suppl 3):280. <https://doi.org/10.1186/s13063-017-1978-4>.
63. Kirkham JJ, Davis K, Altman DG, et al. Core Outcome Set-STAndards for Development: The COS-STAD recommendations. *PLoS Med*. 2017;14(11):e1002447. <https://doi.org/10.1371/journal.pmed.1002447>.
64. Kohavi R, Tang D, Xu Y, et al. Online randomized controlled experiments at scale: lessons and extensions to medicine. *Trials*. 2020;21(1):150. <https://doi.org/10.1186/s13063-020-4084-y>.
65. Radić M, Brinkmann C, Radić D, et al. Digitale Gesundheitsanwendungen auf Rezept: Wie steht es um die Akzeptanz in der Ärzteschaft 2021. Available at [https://www.imw.fraunhofer.de/content/dam/moez/de/documents/210303\\_Studie\\_Digitale%20Gesundheitsanwendungen%20auf%20Rezept\\_DiGAs.pdf](https://www.imw.fraunhofer.de/content/dam/moez/de/documents/210303_Studie_Digitale%20Gesundheitsanwendungen%20auf%20Rezept_DiGAs.pdf). Accessed 19 Sept 2022.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

