

# Evidence of Digital Health Applications from a State-Regulated Repository for Reimbursable Health Applications in Germany

Urs-Vito ALBRECHT<sup>a,1</sup>, Ute VON JAN<sup>b</sup>, Dennis LAWIN<sup>a,c</sup>,  
Evgenii PUSTOZEROV<sup>a</sup> and Florian DITTRICH<sup>a,d</sup>

<sup>a</sup>*Department of Digital Medicine, Medical Faculty OWL, Bielefeld University, Bielefeld, Germany*

<sup>b</sup>*Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School, Hannover, Germany*

<sup>c</sup>*Department of Cardiology and Intensive Care Medicine, University Hospital OWL of Bielefeld University, Campus Klinikum Bielefeld, Bielefeld, Germany*

<sup>d</sup>*Joint Centre Bergisch Land, Department for Orthopaedics, Sana Fabricius Clinic Remscheid, Remscheid, Germany*

ORCID ID: Urs-Vito Albrecht <https://orcid.org/0000-0002-8989-6696>, Ute von Jan <https://orcid.org/0000-0001-9225-593X>,  
Dennis Lawin <https://orcid.org/0000-0002-1292-7845>,  
Evgenii Pustozerov <https://orcid.org/0000-0001-8109-1319>,  
Florian Dietrich <https://orcid.org/0000-0002-5135-4736>

**Abstract.** 17 RCTs for 15 digital health applications (DiGA) permanently listed in the state-regulated register were analyzed descriptively for methodological study aspects relevant to evidence analysis. The analysis revealed that several underlying studies had limitations, at least worthy of discussion, in terms of their power concerning sample size, intervention and control group specifications, drop-out rates, and blinding.

**Keywords.** health apps, evidence, reimbursement, digital health application

## 1. Introduction

In Germany, the Digital Health Care Act (DVG), passed in November 2019, enables doctors and psychotherapists to prescribe “DiGA” (digital health applications) that are covered for those enrolled in statutory health insurance, based on criteria specified in the DVG [1] and the Digital Health Application Regulation (DiGAV) [2]. DiGA are defined as digital, certified low-risk medical devices that help “[...] to support the detection, monitoring, treatment or alleviation of diseases or the detection, treatment, alleviation or

---

<sup>1</sup> Corresponding Author: Urs-Vito Albrecht, Department of Digital Medicine, Medical Faculty OWL, Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany; E-mail: [urs-vito.albrecht@uni-bielefeld.de](mailto:urs-vito.albrecht@uni-bielefeld.de).

compensation of injuries or disabilities [...]” [1]. This covers not only apps but also browser-based applications. However, only applications listed in the directory of digital health applications (DiGA-Verzeichnis, DiGA-VZ) at the German Federal Office for Drugs and Medical Devices (BfArM) [3], after successful completion of an assessment, can be prescribed by physicians or therapists or reimbursed after approval by the health insurer. Permanent inclusion in this directory occurs only if an application has successfully demonstrated interoperability, adequate consideration of data protection and data security, meets the requirements for medical device status (safety, functionality, quality), and if it has demonstrated a “positive impact on care” (PIC). In the absence of sufficient evidence regarding the positive impact of an app, it is possible to provisionally include the app in the directory for a limited period (12 months, extension to 24 months max. is possible), during which time corresponding evidence has to be provided. In this article, the authors provide descriptive data for permanently listed applications and discuss the acquired data under evidence aspects.

## 2. Method

Available meta-tagged information was extracted from the web-based interface provided by BfArM (DiGA-Verzeichnis, [3]) on December 5th, 2022. The data was stored in tabular form and initially evaluated with respect to the apps’ listing status (i.e., provisional or permanent). Entries with a permanent status were subjected to further descriptive analysis (counts, percentages (%), mean values (m), standard deviations (sd), median (md), interquartile range (iqr)) with respect to indication, primary endpoints, control group, and setting, the number of participants, percentage of drop-outs, duration of use and follow-up within the study context, blinding status, National Institute for Health and Care Excellence (NICE) classification [4], and Agency for Healthcare Research and Quality (AHRQ) criteria [5,6].

## 3. Results

The DiGA directory [3] contained 15 permanently listed applications. Prices ranged between 189€ and 599€ (m: 333,43€; sd: 169,68€). 4 DiGA were available natively for iOS and Android, and there were 8 purely web-based apps and 3 with both web and native implementations. 9 DiGA were dedicated to the BfArM category “psyche,” 2 DiGA to “hormones and metabolism,” and 1 DiGA each was assigned to the categories “muscles, bones and joints,” “nervous system”, “ears,” and “other”. A listing of the indications addressed by the applications can be found in Table 1. For all permanently listed apps, published evidence has been deemed to sufficiently support medical benefit and efficacy according to the BfArM’s requirements. However, there are not always links to peer-reviewed publications in this regard (see the cited publications and footnotes in Table 1). Also, all DiGA fully meet the corresponding evidence requirements of the NICE category 3b (i.e., therapeutic purposes, meaning the app “provides treatment for a diagnosed condition or guides treatment decisions”), independent of BfArM logic (“Evidence of positive benefit-risk ratio by valid comparative studies or at least one RCT”). 14 of these apps have an AHRQ evidence level of “Ib” (“At least one sufficiently large, methodologically high-quality RCT”), and for one app, there is higher level

evidence at level “Ia” (“At least one meta-analysis based on methodologically high-quality randomized controlled trials”).

**Table 1.** Methodological study aspects of 15 permanently registered DiGA.

No	Primary Endpoint	Ctrl.	Part.[n]	DrpOut I; C[%]	Use	FlwUp[m]	Bld	AHQR
01a <sup>[7]</sup>	Depression	std	1013	26; 25	3	6	yes <sup>a</sup>	Ib
01b <sup>[8]</sup>	Depression	std	163	27; 28	3	6	no	Ib
01c <sup>[9]</sup>	Depression	mixed	2901	– <sup>b</sup> ; – <sup>b</sup>	2–3	Mixed	no <sup>c</sup>	Ia
02 <sup>[10]</sup>	MS related fatigue	std	275	32; 15	3	24	no	Ib
03 <sup>[11]</sup>	depression (diabetics)	online education	260	24; 12	2	– <sup>d</sup>	no	Ib
04 <sup>[12],e</sup>	Panic	wait list	92	22; 21	2	6	yes <sup>f</sup>	Ib
05 <sup>[13]</sup>	Stress	wait list	264	30; 8 <sup>g</sup>	1.75	12	no	Ib
06 <sup>[14],h</sup>	Vaginism	wait list	200	–; – <sup>i</sup>	3	6	no	Ib
07 <sup>–h</sup>	phobia/panic	wait list + support	297	18–21; 6–19 <sup>j</sup>	2	6	no	Ib
08 <sup>[15],h</sup>	Tinnitus	std	187	–; – <sup>k</sup>	3	9/12 <sup>l</sup>	no	Ib
09 <sup>[16],m</sup>	Depression	wait list	401	57-59 <sup>n</sup> ; 71	3 <sup>o</sup>	6	yes <sup>p</sup>	Ib
10 <sup>–h</sup>	Anxiety	wait list	156	27; 13	3	3	no	Ib
11 <sup>[17]</sup>	Insomnia	wait list	56	10; 0	1.5	12 <sup>q</sup>	no	Ib
12 <sup>[18]</sup>	Anxiety	std	139	20; 13	3	6	no	Ib
13 <sup>[19],r</sup>	back pain	std	215	–; 2	10	3	no	Ib
14 <sup>[20]</sup>	alcohol consumption	std	608	37; 23	6	6	no	Ib
15 <sup>[21],s</sup>	obesity	cont. as usual	149	7 <sup>t</sup>	9	12	no	Ib

All 17 corresponding studies were prospective randomized controlled trials (RCT) following the Intention-to-Treat principle. For one application (01), there was also a meta-analysis of multiple RCTs related to the app. Legend: “Ctrl.”: Control group; “Part.[n]”: number of participants; “DrpOut I;C[%]”: Dropout rates for intervention and control group; “Use, FlwUp[m]”: use and follow up periods in months; “Bld”: blinding; std: standard care

<sup>a</sup> Assessor-blinded.

<sup>b</sup> 12 studies were included, with dropout rates between 6 and 56%.

<sup>c</sup> Assessor blinding was impossible, as only studies with self-reported assessments were included.

<sup>d</sup> While the study (available with study id DRKS00004748 on <https://drks.de/>) is described as using a 6 and 12-month follow-up assessment, the publication [11] does not provide data related to this.

<sup>e</sup> Publication listed on <https://hellobetter.de/online-kurse/panik/>.

<sup>f</sup> Interviewers were blinded to the participant’s randomization status.

<sup>g</sup> Data reported at 12 months for intervention (dropout at 6 months: 12.9%), 6 months for the control group.

<sup>h</sup> References to peer-reviewed publications lacking in the DiGA directory, but specified on the DRKS homepage <https://drks.de/> (App 06: DRKS00010228, App 08: DRKS00022973, App 10: DRKS00023799) or on <https://clinicaltrials.gov/>: App 07: NCT0551080).

<sup>i</sup> Dropout data pay-walled; the abstract states: “on average, participants completed 79% of the intervention.”

<sup>j</sup> Values/diagnosis: Agoraphobia, I: 18%, C: 6%; panic dis., I: 21%, C: 12%; social phob., I: 21%, C: 19%.

<sup>k</sup> Dropout values not stated within the provided data [15].

<sup>l</sup> 9 months for intervention, 12 months for the control group.

<sup>m</sup> Publication identified from the manufacturer’s homepage.

<sup>n</sup> Dropout for follow-up, two intervention groups (guided: 57%, unguided: 59%).

<sup>o</sup> Intervention was available after the initial 12 weeks until follow-up.

<sup>p</sup> “Interviewers were blinded to the assigned group of individuals” [16].

<sup>q</sup> Intervention group only.

<sup>r</sup> No peer-reviewed publication for the RCT is listed in the directory, however, recently stated on the manufacturer’s homepage.

<sup>s</sup> The preliminary RCT data (not yet published in a peer-reviewed journal) not specified in the DiGA directory, but identified via PubMed, describes the use by approx. 11,000 users in Germany. Data shown here were, however, obtained from the DiGA directory.

<sup>t</sup> Value for intervention and control group combined (at 9 months).

#### 4. Discussion

In addition to fulfilling all technical requirements of the DiGA, it is also crucial to evaluate the medical benefit by providing evidence for PIC. By creating this new definition, the BfArM has given its concept of quality a framework in terms of terminology. However, the prerequisite for this is submitting a scientific evaluation concept prepared by a “manufacturer-independent institution to prove PIC, as well as the medical services required for testing” [1]. All 15 permanently listed applications in the German DiGA repository show at least an evidence level of Ib (AHRQ) and a NICE category of 3b. However, there are methodological limitations among the underlying RCTs to some degree, and thus they carry a higher risk of bias. For example, the number of study participants in the RTCs is quite small compared to numbers commonly included in clinical trials (md: 215; iqr: 141), which raises questions about the validity of the studies. The frequently higher dropout rates in the intervention groups (md: 26%; iqr: 10%) compared with the control groups (md: 13%; iqr: 15%) may indicate problems with the internal validity of some RCTs. Possible systematic errors may be responsible for these dropout rates. In most cases, the definition of the “standard of care” (SOC) applied by the investigators is ambiguous. Therefore, the interventions' advantages over SOC are not amenable to interpretation. Overall, the study endpoints are often only imprecisely aligned with the intervention objectives. Also, it is often unclear whether participants assigned to the control groups were given any previous treatment or were naive to treatment before randomization. Another factor for bias is the lack of blinding against interventions: only 3 of the 17 publications we evaluated were blinded. Quality control is debatable for 3 studies where we could not find information about external peer reviews, as they appeared to be published exclusively on the manufacturer’s homepage. The greater question of external validity could not be answered with designs that cover only a few months of usage (md: 3m; IQR: 1m) or short follow-ups (md: 6m; IQR: 6m). Only evaluations with a longer-term application under everyday conditions in post-marketing studies could help to address this aspect properly.

#### 5. Conclusions

All 15 applications permanently listed in the German DiGA repository provided evidence supported by RCTs. However, on closer inspection, some of the underlying studies have at least debatable limitations regarding their validity. The quality of the studies already available, nevertheless, exceeds the low minimum requirements for inclusion in the repository under federal law. Nonetheless, further studies are needed to improve the strength of evidence on the benefits of apps on prescription. Evidence gaps need to be closed. This is not only true for Germany, but for other countries that have either already implemented or are in the process of designing policies to be applied in the context of health apps [22].

#### References

- [1] Deutscher Bundestag. Digitale-Versorgung-Gesetz – DVG. Fassung vom 08.11.2019, verabsch. im Bundesrat am 29.11.2019; 2019. [cited 2022 Dec 9]. Available from: <https://www.bundesrat.de/SharedDocs/drucksachen/2019/0501-0600/557-19.pdf>.

- [2] Bundesamt für Justiz. Digitale Gesundheitsanwendungen-Verordnung – DiGAV; 2020. [cited 2023 Jan 1]. Available from: <https://www.gesetze-iminternet.de/digav/BJNR076800020.html>.
- [3] BfArM. DiGA-Verzeichnis; 2022. [cited 2022 Dec 9]. Available from: <https://diga.bfarm.de/de/verzeichnis>.
- [4] NICE. Evidence Standards Framework for Digital Health Technologies. NHS England; 2019. Available from: <https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/evidence-standards-framework/digital-evidence-standards-framework.pdf>.
- [5] Steiner S, Lauterbach KW. Development of evidence-based clinical practice guidelines; a model project integrating external evidence and clinical expertise. *Med Klin (Munich)*. 1999 Nov;94(11):643-7.
- [6] Berkman ND, Lohr KN, Ansari M, McDonagh M, Balk E, Whitlock E, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. In: *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville (MD): Agency for Healthcare Research and Quality (US), <https://www.ncbi.nlm.nih.gov/pubmed/24404627> (2013).
- [7] Klein JP, Berger T, Schroder J, Späth C, Meyer B, Caspar F, 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 May;85(4):218-28.
- [8] Meyer B, Bierbrodt J, Schroder J, Berger T, Beevers CG, Weiss M, et al. Effects of an Internet intervention (Deprexis) on severe depression symptoms: randomized controlled trial. *Internet Interventions*. 2015;2(1):48-59.
- [9] Twomey C, O'Reilly G, Bultmann O, Meyer B. Effectiveness of a tailored, integrative Internet intervention (deprexis) for depression: Updated meta-analysis. *PLoS One*. 2020 Jan;15(1):e0228100.
- [10] Pöttgen J, Moss-Morris R, Wendebourg JM, Feddersen L, Lau S, Köpke S, et al. Randomised controlled trial of a self-guided online fatigue intervention in multiple sclerosis. *J Neurol Neurosurg Psychiatry*. 2018 Sep;89(9):970-6.
- [11] Nobis S, Lehr D, Ebert DD, Baumeister H, Snoek F, Riper H, 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 May;38(5):776-83.
- [12] Ebenfeld L, Lehr D, Ebert DD, Kleine Stegemann S, Riper H, Funk B, et al. Evaluating a Hybrid Web-Based Training Program for Panic Disorder and Agoraphobia: Randomized Controlled Trial. *J Med Internet Res*. 2021 Mar;23(3):e20829.
- [13] Heber E, Lehr D, Ebert DD, Berking M, Riper H. Web-based and mobile stress management intervention for employees: A randomized controlled trial. *J Med Internet Res*. 2016 Jan;18(1):e21.
- [14] Zarski AC, 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 Nov;89(11):909-24.
- [15] Kalmeda. Tinnitus-Studie; 2022. [cited 2022 Dec 10]. Available from: <https://www.kalmeda.de/fachkreise/tinnitus-studie>.
- [16] Kramer R, Köhne-Volland L, Schumacher A, Köhler S. Efficacy of a Web-Based Intervention for Depressive Disorders: Three-Arm Randomized Controlled Trial Comparing Guided and Unguided Self-Help With Waitlist Control. *JMIR Form Res*. 2022 Apr;6(4):e34330.
- [17] Lorenz N, Heim E, Roetger A, Birrer E, Maercker A. Randomized controlled trial to test the efficacy of an unguided online intervention with automated feedback for the treatment of insomnia. *Behav Cogn Psychother*. 2019 May;47(3):287-302.
- [18] Berger T, Urech A, Krieger T, Stolz T, Schulz A, Vincent A, 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 Jan;47(1):67-80.
- [19] Weise H, Zenner B, Schmiedchen B, Benning L, Bulitta M, Schmitz D, et al. The Effect of an AppBased Home Exercise Program on Self-reported Pain Intensity in Unspecific and Degenerative Back Pain: Pragmatic Open-label Randomized Controlled Trial. *J Med Internet Res*. 2022 Oct;24(10):e41899.
- [20] Zill JM, Christalle E, Meyer B, Harter M, Dirmaier J. The effectiveness of an internet intervention aimed at reducing alcohol consumption in adults. *Dtsch Arztebl Int*. 2019 Feb;116(8):127-33.
- [21] Forkmann K, Roth L, Mehl N. Introducing zanadio – A Digitalized, Multimodal Program to Treat Obesity. *Nutrients*. 2022 Aug;14(15).
- [22] Essén A, Stern AD, Haase CB, Car J, Greaves F et al. Health app policy: international comparison of nine countries' approaches. *NPJ Digit Med* 2022; 5: 31.