Systematic evaluation of content and quality of English and German pain apps in European app stores

Yannik Terhorst [a,b,*], Eva-Maria Messner [b], Dana Schultchen [c], Sarah Paganini [d], Alexandra Portenhauser [b], Anna-Sophia Eder [b], Melanie Bauer [c], Mike Papenhoff [e], Harald Baumeister [b], Lasse Bosse Sander [f]

[a] Department of Research Methods, Institute of Psychology and Education, Ulm University, Albert-Einstein-Allee 47, 89069 Ulm, Germany
[b] Department of Clinical Psychology and Psychotherapy, Institute of Psychology and Education, Ulm University, Lise-Meitner-Str. 16, 89081 Ulm, Germany
[c] Department of Clinical and Health Psychology, Institute of Psychology and Education, Ulm University, Albert-Einstein-Allee 47, 89069 Ulm, Germany
[d] Department of Sport Psychology, University of Freiburg, Schwarzwaldstraße 175, 79117 Freiburg, Germany
[e] Pain Medicine Clinic, BG Klinikum Duisburg, Grossenbaumer Allee 250, D-47249 Duisburg, Germany
[f] Department of Rehabilitation Psychology and Psychotherapy, Institute of Psychology, Albert-Ludwigs-University Freiburg, Engelberger Str. 41, 79106 Freiburg im Breisgau, Germany

ABSTRACT

Background and objective: Pain spans a broad spectrum of diseases and types that are highly prevalent and cause substantial disease burden for individuals and society. Up to 40% of people affected by pain receive no or inadequate treatment. Providing a scalable, time-, and location-independent way for pain diagnostic, management, prevention and treatment mobile health applications (MHA) might be a promising approach to improve health care for pain. However, the commercial app market is rapidly growing and unregulated, resulting in an opaque market. Studies investigating the content, privacy and security features, quality and scientific evidence of the available apps are highly needed, to guide patients and clinicians to high quality MHA.
Contributing to this challenge, the present study investigates the content, quality, and privacy features of pain apps available in the European app stores.
Methods: An automated search engine was used to identify pain apps in the European Google Play and Apple App store. Pain apps were screened and checked for systematic criteria (pain-relatedness, functionality, availability, independent usability, English or German). Content, quality and privacy features were assessed by two independent reviewers using the German Mobile Application Rating Scale (MARS-G). The MARS-G assesses quality on four objectives (engagement, functionality, aesthetics, information quality) and two subjective scales (perceived impact, subjective quality).
Results: Out of 1034 identified pain apps 218 were included. Pain apps covered eight different pain types. Content included basic information, advice, assessment and tracking, and stand-alone interventions. The overall quality of the pain apps was average M = 3.13 (SD = 0.56, min = 1, max = 4.69). The effectiveness of less than 1% of the included pain apps was evaluated in a randomized controlled trial. Major problems with data privacy were present: 59% provided no imprint, 70% had no visible privacy policy.
Conclusion: A multitude of pain apps is available. Most MHA lack scientific evaluation and have serious privacy issues, posing a potential threat to users. Further research on evidence and improvements privacy and security are needed. Overall, the potential of pain apps is not exploited.

1. Introduction

Pain is a highly prevalent and manifold condition (Breivik et al., 2006; Jackson et al., 2016; Leadley et al., 2012). According to the global burden of disease study, pain is the number one cause for years lived with disability (YLD) (James et al., 2018). Low back pain alone caused

65 Million YLDs in the year 2017 (James et al., 2018), and is associated with high societal costs (Baumeister et al., 2012; Breivik et al., 2013; Gaskin and Richard, 2012).

Multiple effective treatment options for pain exist (e.g. pharmacological, interventional, physiological or psychological treatments) (Turk et al., 2011; Veehof et al., 2016). Pain severity often results from a complex interaction between 1) pathological mechanisms, 2) past pain experiences and social environment, 3) cognitive factors and 4) emotional factors (Turk et al., 2011). Thus, an interdisciplinary and multidimensional treatment approach has been established as the gold standard (Kumar, 2007; Traue et al., 2013). However, 40% of affected individuals do not receive adequate treatment and are unsatisfied with the effectiveness of their treatment of long-lasting pain (Breivik et al., 2006). Insufficient availability and involvement of pain specialist, inadequate medication, infrastructural barriers (e.g., traveling time), as well as personal barriers (e.g., limited time resources) contribute to insufficient health care (Becker et al., 2017; Breivik et al., 2006, 2013; Rod, 2016).

Given their time- and location-independent nature, mobile health applications (MHA) might provide an opportunity to improve health care for people with pain (Albrecht, 2016; Donker et al., 2013; Linardon et al., 2019; Thurnheer et al., 2018; Weisel et al., 2019). Pain apps can be used in various ways to improve the current healthcare situation, including 1) diagnostics, 2) screening, 3) self-monitoring, 4) referral, 5) education, 6) pain-management, and 7) social support (Albrecht, 2016; Lalloo et al., 2015, 2017), accompanying on-site treatments or being used as standalone interventions (Ebert et al., 2017, 2018; Messner et al., 2019a). This might be in particular a valuable addition to healthcare for pain during times of lockdown, social distancing, and other quarantine measurements that affect pain treatment and management (El-Tallawy et al., 2020). However, besides their structural advantages MHA and pain apps can also offer anonymous health care to individuals who fear stigmatization that may not seek healthcare otherwise (Andrade et al., 2014).

The commercial app market is rapidly growing and multiple pain apps are available (Lalloo et al., 2015; Salazar et al., 2018; Wallace and Dhingra, 2014) and other health conditions (e.g. depression, anxiety, rheumatism, PTSD: (Knitza et al., 2019; Sander et al., 2019; Sucala et al., 2017; Terhorst et al., 2018, 2020)). In contrast to the multitude of apps available in the app stores, there commonly is no or only limited evidence for their effectiveness (Donker et al., 2013; Rathner and Probst, 2018; Sucala et al., 2017; Terhorst et al., 2018; Thurnheer et al., 2018). Besides the lack of evidence, several studies have highlighted issues with privacy policies (e.g., missing or insufficient) (Grundy et al., 2019; Huckvale et al., 2019), lack of health care professional involvement (Lalloo et al., 2015; Wallace and Dhingra, 2014) or negative side-effects in case of device failure (Luxton et al., 2011).

Another concern is the limited availability of information about MHA in the commercial app stores: App description, user star ratings, and user comments are often the only, but an unreliable source for users to assess MHA content and quality (Armstrong, 2015; Shen et al., 2015; Stoyanov et al., 2015, 2016; Terhorst et al., 2018): For users, perceived MHA quality is rather led by perception of MHA look and feel than on information and content quality, which are paramount from clinicians perspective (Armstrong, 2015; Bardus et al., 2016; Nicholas et al., 2015; Stoyanov et al., 2016; Terhorst et al., 2018). Consequently, users and health care providers are experiencing difficulties in identifying MHA suiting their conditions and needs (Shen et al., 2015). Hence, studies investigating the content, privacy and security features, quality and scientific evidence of the available apps are highly needed, to guide patients and clinicians to high quality MHA (Donker et al., 2013; Knitza et al., 2019; Sander et al., 2020; Schoeppe et al., 2017; Shen et al., 2015; Terhorst et al., 2018).

In prior investigations Wallace and Dhingra (2014), and Lalloo et al. (2015) provided a comprehensive overview of the content of pain apps. However, they did not review the quality of MHA (Lalloo et al., 2015;

Wallace and Dhingra, 2014). Extending the content analysis, Salazar et al. (2018) provided an overview on quality and content of 18 pain apps limited to MHA in the Spanish app stores. Overall, they reported average quality of the included MHA with the lowest quality in the dimensions engagement and information quality. However, next to the Spanish focus, their search term was limited to generic search term "pain" (both in English and Spanish), which seems too narrow given the broad spectrum of pain disorders. This limitation is also true for the previously mentioned reviews by Wallace and Dhingra (2014) (search term: "pain"), and Lalloo et al. (2015) (search term: "pain" and "pain management").

Hence, this study aims to extend the previous findings by an improved search, a more detailed evaluation of privacy and content features, reviewing the quality of pain apps, and investigating the relationship between standardized expert ratings of quality and user-star ratings.

The following questions will be addressed:

1. Which content is provided by pain apps in the European commercial app stores?
2. What measures have been taken to ensure privacy and data protection?
3. Of which quality are pain apps in the European commercial app stores regarding user engagement, functionality, aesthetics, and information content?
4. Is the user star rating associated with the standardized MARS-G quality rating?

2. Methods

2.1. Search strategy and inclusion criteria

An automatic search engine (webcrawler) developed in context of the mobile health app database project (MHAD; http://mhad.science/) (Messner et al., 2021; Meßner et al., 2019; Paganini et al., 2019; Portenhauser et al., 2021; Sander et al., 2019; Schultchen et al., 2020; Terhorst et al., 2018) was used to systematically search the European Google Play and Apple App store for pain apps. The webcrawler searches and automatically extracts the available information from the app stores. Search terms included were: "chronic pain", "pain", "ache", "migraine", "headache", "rheumatoid arthritis", "back pain" and "pain diary". All identified apps were listed in a central database, and duplicates were automatically removed. The search was conducted on October 24th, 2018.

After identification, a two-step procedure was applied to assess the eligibility of identified apps: In a first step, titles and descriptions were screened by one reviewer (either DS, LS, MB, YT). Inclusion criteria were: (a) the app was developed for a pain-related context, (b) the app was available in a language spoken by the reviewer team (English, German), (c) the app was available for download, (d) the app was usable independently (e.g. access not restricted to study participation or a specific clinic). In a second step, all eligible apps were downloaded and checked regarding the criteria on basis of the actual app content. Not working (e.g., unable to start on the device) pain apps were excluded.

2.2. App characteristics

The Mobile Application Rating Scale (MARS) (Messner et al., 2019b; Stoyanov et al., 2015) was used to capture app characteristics: (a) app name, (b) store link (c) platform (android or iOS), (d) content-related subcategory, (e) aims, (f) price, (g) user rating.

2.3. Content and functions

The classification site of the MARS was used to assess content and functions of included pain apps (Messner et al., 2019b; Stoyanov et al.,

2015).

2.4. Privacy and security features

The MARS comes with an assessment of privacy and security features on a descriptive level. All features were assessed based on the downloaded pain apps. External information (e.g. developer web-site) were not investigated, if not referenced within the app.

2.5. Quality rating

The Mobile Application Rating Scale – German (MARS-G), a translated version of the original MARS (Stoyanov et al., 2015) was used for the quality rating (Messner et al., 2019b). The MARS shows good to excellent between rater agreement, reliability, construct validity, and concurrent validity (Messner et al., 2019b; Stoyanov et al., 2015; Terhorst et al., 2020).

Quality was rated on a 5-point scale (1-inadequate, 2-poor, 3-acceptable, 4-good, and 5-excellent). The MARS-G includes 19 items divided into four sub-dimensions: (A) engagement (5 items: fun, interest, individual adaptability, interactivity, target group), (B) functionality (4 items: performance, usability, navigation, gestural design), (C) aesthetics (3 items: layout, graphics, visual appeal), and (D) information quality (7 items: accuracy of app description, goals, quality of information, quantity of information, quality of visual information, credibility, evidence base). Item 19 on the information subscale evaluates the evidence base of a given app ("Has the app been trialled/tested; must be verified by evidence (in published scientific literature)?"). Studies providing evidence for an app were identified by app description, developer's and provider's web-sites, and searching the app name in google scholar and PubMed. Mean scores were calculated for each dimension separately. The overall score represents the mean of the engagement, functionality, aesthetics and information quality scales.

In addition to the objective quality scales of the MARS, the MARS includes two subjective scales: (E) subjective quality (4 items: recommendation, frequency of use, willingness to pay, overall star-rating) and (F) perceived impact (6 items: awareness, knowledge, attitudes, intention to change, help-seeking, behavioural change).

2.6. Rater training and agreement

Psychologists (at least BSc level) conducted the quality assessment. Throughout the assessment, all raters were supervised by DS or LS (licensed psychotherapist). Both supervisors had experience with using the MARS (Messner et al., 2021; Portenhauser et al., 2021; Sander et al., 2019; Schultchen et al., 2020; Terhorst et al., 2018). All raters underwent standardized online training before the quality assessment (https://youtu.be/5vwMiCWC0Sc, last updated on January 2021). All pain apps were rated by two independent raters (MB, SP, AP, ASE, DS, YT, LS). Agreement between raters was evaluated by intra-class correlation (ICC) (Koo and Li, 2016). A cut-off of ICC above 0.75 was defined as the criteria for a satisfactory rater agreement (Fleiss, 1999).

2.7. Associations between user star rating and quality

Associations between MARS ratings and user star-ratings were investigated by means of correlations.

2.8. Data analyses

Rater agreement was examined by intra-class correlation based on a two-way mixed-effects model (Koo and Li, 2016). Consistency between raters was evaluated.

For correlations analysis an alpha-level of 5% was defined. P-values were adjusted using the procedure proposed by Holm (1979).

3. Results

3.1. Search and app characteristics

1034 apps were identified (Apple App Store = 693, Play Store = 341). After screening and eligibility check, a total of 218 pain apps were eligible for inclusion (Apple App Store = 110, Play Store = 108) (see Fig. 1).

Of all 218 included pain apps, 108 pain apps had a user-star-rating (M = 4.05; SD = 0.69), while 110 were not rated by users. The majority of pain apps were free of charge ($n = 170$, 78%). Prices of fee-based pain apps ranged from 1.09 EUR to 21.99 EUR (M = 3.88 EUR, SD = 3.17). Pain apps were identified in eight store categories (multiple categories can be assigned to an app): "health and fitness" ($n = 145$, 67%), "medical" ($n = 125$, 57%), "lifestyle" ($n = 14$, 6%), "education" ($n = 8$, 4%), "books and reference" ($n = 5$, 2%), "news and magazines" ($n = 3$, 1%), "sport" ($n = 1$, 1%), "entertainment" ($n = 1$, 1%), "utilities" ($n = 1$, 1%).

The included pain apps target a broad range of areas. Most pain apps focused on headache and migraine ($n = 59$, 27%), back pain ($n = 53$, 24%), arthritis and rheumatism ($n = 30$, 14%), chronic pain ($n = 18$, 8%), general/unspecified pain ($n = 18$, 8%), fibromyalgia ($n = 3$, 1%), neck ($n = 2$, 1%) and other specific pain types ($n = 4$, 2%). Pain apps offering a pain assessment or diary function without focus on a specific pain type were grouped under "Diary & Assessment" ($n = 31$, 14%).

Overall most of the pain apps aimed for improved physical fitness ($n = 151$, 69%), general well-being ($n = 130$, 60%), goal-setting ($n = 35$, 16%), behavior change ($n = 22$, 10%), reduction of stress ($n = 9$, 4%), emotion regulation ($n = 5$, 2%), social-behavior ($n = 5$, 2%), reduction of anxiety ($n = 4$, 2%), entertainment ($n = 4$, 2%), reduction of depression ($n = 3$, 1%), addiction ($n = 1$, <1%) and other ($n = 58$, 27%). For a more nuanced representation of the aims in respect to the type of pain (e.g., back pain, headache and migraine) see Multimedia Appendix 1.

3.2. Content and functions

Almost half ($N = 106$, 49%) of MHA provided information about pain, 47% ($N = 102$) offered advice, 44% ($N = 95$) included a tracking feature, and 29% ($N = 64$) included an assessment tool. Furthermore, frequent features were physical exercise ($N = 59$, 27%), feedback ($N = 40$, 18%), strategies ($N = 36$, 17%), and reminder ($N = 35$, 16%). An overview of the content and functions grouped by the type of pain is presented in Multimedia Appendix 1.

3.3. Privacy and security features

Of all included pain apps, $n = 76$ (35%) had a legal disclosure, $n = 55$ (25%) a visible privacy policy, $n = 32$ (15%) required a log-in, $n = 33$ (15%) offered password protection, $n = 21$ (10%) informed about conflict of interests/financial background. Consent to data collection was required in an active form by $n = 26$ pain apps (12%) and in a passive form by 20 pain apps (9%). Of the pain apps including assessment ($n = 56$), active consent was required by 9 (16%), and passive consent was present in 13 (23%) pain apps.

3.4. Quality rating

Agreement between raters was excellent (ICC = 0.84, 95%-CI: 0.83 to 0.85). The overall quality of pain apps was average M = 3.13 (SD = 0.56), ranging from 1.00 to 4.69. Quality ratings of the four sub-dimensions yielded: functionality M = 3.66 (SD = 0.62), aesthetics M = 3.11 (SD = 0.73), information quality M = 2.93 (SD = 0.69) and engagement M = 2.81 (SD = 0.75). The quality per type of pain is summarized in Table 1. A more detailed summary of pain apps with at least good overall quality (>4.0) is presented in Table 2. For the quality
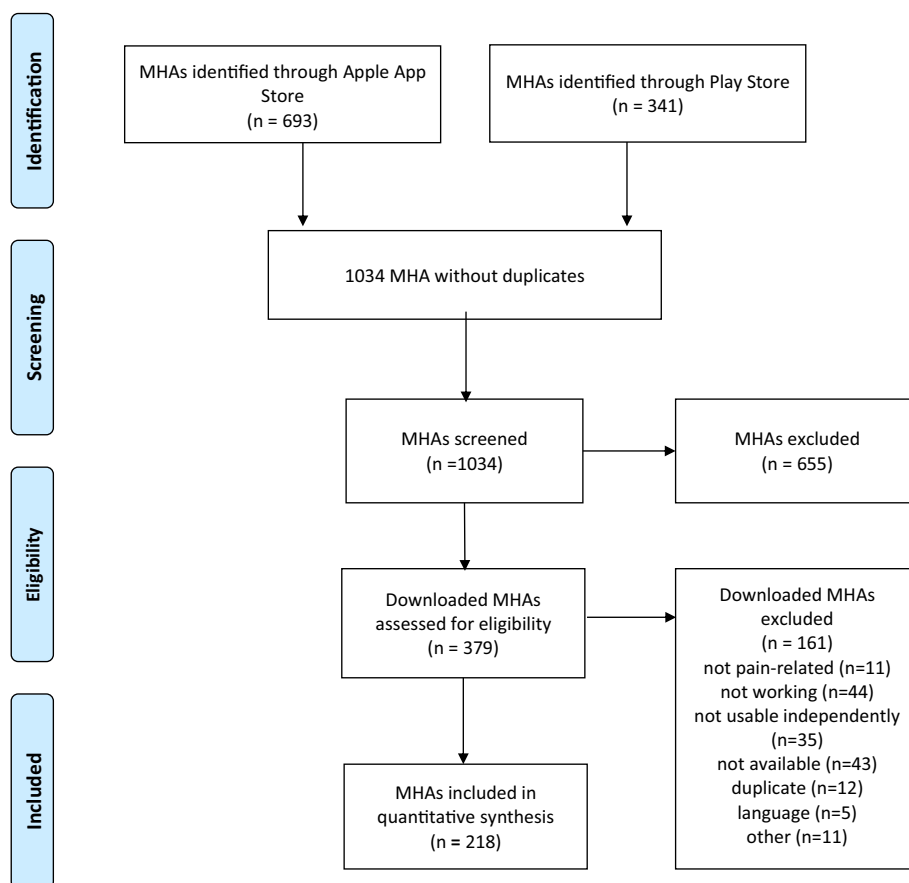
Fig. 1. Flow-chart of the in- and exclusion process.

ratings of all pain apps, their content and privacy features see Multi-media Appendix 1.

Besides the four objective dimension the MARS includes a scale for the subjective quality and the perceived impact on health. Across all pain apps the trained reviewers rated subjective quality M = 2.19 (SD = 0.70) and the perceived impact on health M = 2.06 (SD = 0.69) as poor. See Multimedia Appendix 1 for the app-wise ratings.

3.5. Quality rating on evidence

The evidence item of the information quality section indicated that only a single pain app (="Kaia: Back Pain Relief at Home") was tested in a randomized controlled trial (Toelle et al., 2019) and the same pain app was also tested in a cluster randomized controlled trial (cRCT) (Priebe et al., 2020) showing its superiority compared to the control conditions. Overall this translated to an RCT evidence for 0.5% of all examined pain apps. Pain apps rated with at least 3 (="App has been trialled (e.g., acceptability, usability, satisfaction ratings) and has positive outcomes

in studies that are not RCTs, and there is no contradictory evidence."; (Stoyanov et al., 2015)) were the English "GeoPain" (https://www.geopain.com/) and the German "Migräne App" (https://schmerzklinik.de/die-migraene-app/). The evidence item of a total of 199 pain apps (91.3%) were rated as "not applicable", since no information on evidence were found.

3.6. Associations between user-rating and quality

User ratings were available for 108 pain apps (49.5%). No significant correlations between user star-ratings and standardized expert ratings were found ($p > .05$; overall $r = 0.03$, engagement $r = 0.04$, functionality $r = -0.08$, aesthetics $r = 0.01$ and information quality $r = 0.00$).

4. Discussion

This study investigated the content, privacy and security features, quality of pain apps. Identified pain apps cover a broad range of pain

Table 1
Quality of MHA.

Area of application	Frequency	Engagement	Functionality	Aesthetics	Information	Overall
Arthritis & rheumatism	30	2.97 (0.77)	3.70 (0.63)	3.18 (0.78)	3.39 (0.74)	3.31 (0.60)
Back pain	53	2.44 (0.71)	3.44 (0.61)	2.77 (0.83)	2.71 (0.63)	2.84 (0.58)
Chronic pain	18	2.88 (0.74)	3.65 (0.51)	3.04 (0.86)	2.83 (0.61)	3.10 (0.55)
Diary & assessment	31	3.06 (0.63)	3.69 (0.55)	3.36 (0.53)	2.91 (0.61)	3.25 (0.48)
Fibromyalgia	3	2.57 (0.25)	3.92 (0.62)	2.89 (0.75)	2.96 (1.04)	3.08 (0.62)
General/unspecific	18	2.94 (0.79)	3.70 (0.57)	3.28 (0.50)	3.29 (0.68)	3.30 (0.46)
Headache & migraine	59	2.92 (0.73)	3.87 (0.56)	3.28 (0.60)	2.87 (0.64)	3.24 (0.45)
Neck	2	2.50 (0.14)	4.06 (0.09)	3.33 (0.00)	2.92 (0.35)	3.20 (0.15)
Other	4	2.58 (1.05)	2.66 (1.26)	2.46 (1.01)	2.50 (1.08)	2.55 (1.05)
Total	218 (100%)	2.81 (0.75)	3.66 (0.62)	3.11 (0.73)	2.93 (0.69)	3.13 (0.56)

Table 2
High-quality MHA.

App	Pain area	Content	Engagement	Functionality	Aesthetics	Information	Overall	Subjective quality	Perceived impact
Jointfully Osteoarthritis	Arthritis & rheuma	Assessment, tracking, feedback, information, advice, goalsetting, strategies, resources, physical, reminder, tailored, medicine	4.30	4.00	4.33	4.43	4.27	3.63	4.08
RheumaLive	Arthritis & rheuma	Tracking, feedback	3.60	4.75	4.00	4.33	4.17	3.25	1.42
Kaia: Back Pain Relief at Home	Back pain	Tracking, information, advice, strategies, alternative medicine, physical	4.30	4.50	4.67	3.71	4.30	3.63	3.67
Schmerztagebuch [eng.: pain diary] & Community CatchMyPain	Diary & assessment	Assessment, tracking, information, advice,	3.90	4.75	4.17	4.38	4.30	3.63	3.50
Schmerztagebuch [eng.: pain diary] - Pain Tracer	Diary & assessment	Assessment, tracking	3.60	4.50	4.17	4.17	4.11	3.50	2.08
Injurymap	General	Feedback, information, advice, goalsetting, strategies, physical	4.90	4.63	4.50	4.75	4.69	3.50	3.67
Migraine Xplained	Headache & migraine	Assessment, tracking, information, reminder	4.10	5.00	4.67	4.17	4.48	3.75	3.58
Migräne App (Play Store) [eng.:]	Headache & migraine	Tracking, feedback, information, advice, reminder, medicine	4.20	4.50	4.50	4.17	4.34	3.63	3.08
Migräne App (App Store)	Headache & migraine	Assessment, tracking, feedback, information, advice, goalsetting, relaxation	4.40	4.63	4.17	4.13	4.33	3.63	4.00
Migraine Buddy	Headache & migraine	Tracking, feedback	4.60	4.25	4.00	3.21	4.02	3.25	3.25

disorders (e.g. back pain, migraine). Besides the application areas, the content and functions also varied: Most often, they provide information and advice as well as some form of assessment and tracking. Moreover, there were multiple apps that included treatment elements (e.g., relaxation as well as physical, mindfulness, or breathing exercises). Overall the quality of pain apps was acceptable. However, we could only find one RCT and one cRCT evaluating one of the identified pain apps (Priebe et al., 2020; Toelle et al., 2019). Furthermore, for the vast majority (91%) could not find any evidence on the effectiveness and safety of the examined pain apps. The lack of scientific evidence is in line with previous studies in the field of app evaluation (Donker et al., 2013; Meßner et al., 2019; Sander et al., 2019; Sucala et al., 2017; Terhorst et al., 2018, 2020).

Besides the need for scientific evidence, the present findings indicate that the quality of available pain apps could be improved by increasing the engagement of users. Engagement of users reflected the dimension with the weakest score across all subscales. The lack of user engagement is a common problem in MHA across different disorders (Meßner et al., 2019; Sander et al., 2019; Terhorst et al., 2018). Gamification, tracking, and feedback to the users conceptualized as persuasive design features could be promising measurements to overcome this problem (Bakker et al., 2016; Baumeister et al., 2019; Baumel and Yom-Tov, 2018). Moreover, many pain apps would benefit from an improved quality of the provided information. Wrong or misleading information in MHA are a threat to users' safety (Albrecht, 2016; Terhorst et al., 2018). Only if MHA provide evidence-based and high-quality information, they could become a helpful component in health care (Albrecht, 2016; Donker et al., 2013; Terhorst et al., 2018).

In this study, we did not find a significant correlation between user-rating and the MARS-G. Hence, the user-star-rating cannot be considered a reliable measurement to determine the quality of apps. This finding is in line with prior investigations and underlines the need for systematic evaluations guiding users and health care providers to MHA fitting their needs and requirements (Armstrong, 2015; Bardus et al., 2016; Nicholas et al., 2015; Shen et al., 2015; Terhorst et al., 2018).

Since available information in the app stores (e.g., user rating, app description) are insufficient for a reliable assessment of MHA content and quality, other objective platforms to inform users as well as health care providers are needed. Systematic evaluations using reliable and objective instruments to assess content and quality (e.g. MARS (Stoyanov et al., 2015), ENLIGHT (Baumel et al., 2017), APA App Evaluation Model [https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model]) system could function as the foundation of such platforms. Considering the amount of existing reviews, which cover a broad range of health conditions, the building of an objective research-based international information platform for MHA might already be feasible (Bardus et al., 2016; Grainger et al., 2017; Knitza et al., 2019; Machado et al., 2016; Mani et al., 2015; Masterson Creber et al., 2016; Meßner et al., 2019; Salazar et al., 2018; Sander et al., 2019; Santo et al., 2016; Schoeppe et al., 2017; Terhorst et al., 2018; Thornton et al., 2017). However, no international information platform exists and only country specific projects have started to provide evidence-based information such as the German Mobile Health App Database (MHAD; http://mhad.science/), the German HealthOn platform (https://www.healthon.de/) or the US platform PsyberGuide (https://psyberguide.org/).

Equally important to adequate information about MHA quality and content is information about privacy and security (Grundy et al., 2019; Huckvale et al., 2019; Torous et al., 2019). According to international standards and the European Data Protection Regulation MHA should provide a visible privacy policy, an imprint with contact information, and an opt-in consent for all data collection (THE EUROPEAN PARLIAMENT and THE COUNCIL OF THE EUROPEAN UNION, 2018; Torous et al., 2019). Furthermore, users have to be informed about their

rights (e.g., right to inspect collected personal data, right to delete personal data, right to change personal information). However, for most MHA, those criteria were not met. Before MHA become a part of routine healthcare, these criteria should be met to ensure patient safety.

Regarding privacy and security, it has to be kept in mind, that we only assessed the privacy and data security on a descriptive level. In a recent assessment of data sharing and privacy practices of MHA for depression and smoking cessation, a more elaborated procedure was applied, showing not only the lack of privacy policy, but also inadequate or insufficient information in present privacy policies (Huckvale et al., 2019). Future studies should build on this procedure and have a closer look at the privacy policies of MHA not only on a descriptive level, but actively monitoring the data traffic of MHA as well as trying to break or interfere with the data traffic to investigate if data protection criteria are met under attack.

Besides the only descriptive analysis of privacy issues, some further limitations have to be considered. Firstly, this study might have missed relevant MHA due to the restriction to the European Google Play and Apple App store. However, Google Play and Apple App store cover about 99% of the total market, and the loss of relevant pain apps due to the selection of stores should be minimal (StatCounter, 2019). However, it cannot be guaranteed that all included pain apps are also available in the US stores and other major markets (e.g., India, China) or that the present results are transferable to the respective app versions in those stores. Secondly, the search results are limited to the selected search terms and the language restriction (English and German). The terms were selected by experts to cover a broad range of pain-related diseases, symptoms or content. Using this approach this study with $N = 218$ pain apps provides a by far more comprehensive analysis of pain app quality than in a previous review of pain apps ($N = 18$) (Salazar et al., 2018), nonetheless future studies could expand the present findings by including other languages and pain types. Thirdly, while this study makes a valuable contribution to the field by presenting the aims, content and function as well as quality of available pain apps for eight central types of pain (arthritis and rheumatism, back pain, chronic pain, fibromyalgia, headache and migraine, neck, other specific pain types and general pain) and thus, enables patients and health care providers to find a pain app fitting their needs, the study fell short in the comparison of provided content against established treatment guidelines. However, deriving the guidelines from the literature and comparing them with the content of pain apps was not the aim in the present study and doing so in a single study for pain apps in general might not be feasible due to the many pain apps and differences in the guidelines depending on the disorders. Hence, future studies with a closer focus on specific pain disorders (e.g. treatment of back pain) would be more suited to compare pain app content to treatment guidelines.

5. Conclusion

A multitude of pain apps is available. The overall quality of the included pain apps can be considered moderate at best. Particularly, the engagement features and the information quality need to be improved. The vast majority of available pain apps has no evidence and the effectiveness and safety of 99.5% of all examined pain apps have not been proven in an RCT. Above that, issues regarding the lack of privacy policies and information about data security are present. User ratings and app descriptions in the stores are insufficient to inform users. Independent information platforms offering reliable information for users and health care providers about content and quality (e.g., measured by the MARS) are highly needed. So far, the capabilities of pain apps to improve health care are not exploited.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.invent.2021.100376.

References

Albrecht, U., 2016. Chancen Und Risiken von Gesundheits-Apps (CHARISMHA) [Chances and Risks of Mobile Health Applications]. Medizinische Hochschule Hannover.

Andrade, L.H., Alonso, J., Mneimneh, Z., Wells, J.E., Al-Hamzawi, A., Borges, G., Bromet, E., et al., 2014. Barriers to mental health treatment: results from the WHO world mental health (WMH) surveys. Psychol. Med. 44, 1303–1317. https://doi.org/10.1017/S0033291713001943.Barriers.

Armstrong, S., 2015. Which app should I use? BMJ 351, h4597. https://doi.org/10.1136/bmj.h4597.

Bakker, D., Kazantzis, N., Rickwood, D., Rickard, N., 2016. Mental Health Smartphone Apps: Review and Evidence-based Recommendations for Future Developments 3, e7. https://doi.org/10.2196/mental.4984.

Bardus, M., van Beurden, S.B., Smith, J.R., Abraham, C., 2016. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int. J. Behav. Nutr. Phys. Act. 13, 35. https://doi.org/10.1186/s12966-016-0359-9.

Baumeister, H., Knecht, A., Hutter, N., 2012. Direct and indirect costs in persons with chronic back pain and comorbid mental disorders-a systematic review. J. Psychosom. Res. 73 https://doi.org/10.1016/j.jpsychores.2012.05.008.

Baumeister, H., Kraft, R., Baumel, A., Pryss, R., Messner, E.-M., 2019. Persuasive e-health design for behavior change. In: Baumeister, H., Montag, C. (Eds.), Mobile Sensing and Digital Phenotyping: New Developments in Psychoinformatics. Springer, Berlin.

Baumel, A., Yom-Tov, E., 2018. Predicting user adherence to behavioral eHealth interventions in the real world: examining which aspects of intervention design matter most. Transl. Behav. Med. 8, 793–798. https://doi.org/10.1093/tbm/ibx037.

Baumel, A., Faber, K., Mathur, N., Kane, J.M., Muench, F., 2017. Enlight: a comprehensive quality and therapeutic potential evaluation tool for mobile and web-based eHealth interventions. J. Med. Internet Res. 19, e82 https://doi.org/10.2196/jmir.7270.

Becker, W.C., Dorflinger, L., Edmond, S.N., Islam, L., Heapy, A.A., Fraenkel, L., 2017. Barriers and facilitators to use of non-pharmacological treatments in chronic pain. BMC Fam. Pract. 18, 1–8. https://doi.org/10.1186/s12875-017-0608-2.

Breivik, H., Collett, B., Ventafridda, V., Cohen, R., Gallacher, D., 2006. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur. J. Pain 10, 287–333. https://doi.org/10.1016/j.ejpain.2005.06.009.

Breivik, H., Eisenberg, E., O'Brien, T., 2013. The individual and societal burden of chronic pain in Europe: the case for strategic prioritisation and action to improve

knowledge and availability of appropriate care. BMC Public Health 13, 1229. https://doi.org/10.1186/1471-2458-13-1229.

Donker, T., Petrie, K., Proudfoot, J., Clarke, J., Birch, M.-R.R., Christensen, H., 2013. Smartphones for smarter delivery of mental health programs: a systematic review. J. Med. Internet Res. 15, e247 https://doi.org/10.2196/jmir.2791.

Ebert, D.D., Cuijpers, P., Muñoz, R.F., Baumeister, H., 2017. Prevention of mental health disorders using internet- and mobile-based interventions: a narrative review and recommendations for future research. Front. Psychiatry 8, 116. https://doi.org/10.3389/fpsyt.2017.00116.

Ebert, D.D., Van Daele, T., Nordgreen, T., Karekla, M., Compare, A., Zarbo, C., Brugnera, A., et al., 2018. Internet- and mobile-based psychological interventions: applications, efficacy, and potential for improving mental health: a report of the EFPA E-health taskforce. Eur. Psychol. 23, 167–187. https://doi.org/10.1027/1016-9040/a000318.

El-Tallawy, S.N., Nalamasu, R., Pergolizzi, J.V., Gharibo, C., 2020. Pain management during the COVID-19 pandemic. Pain Ther. 9, 453–466. https://doi.org/10.1007/s40122-020-00190-4.

Fleiss, J.L., 1999. The Design and Analysis of Clinical Experiments. John Wiley & Sons, Inc., Hoboken, NJ, USA https://doi.org/10.1002/9781118032923.

Gaskin, D.J., Richard, P., 2012. The economic costs of pain in the United States. J. Pain 13, 715–724. https://doi.org/10.1016/j.jpain.2012.03.009.

Grainger, R., Townsley, H., White, B., Langlotz, T., Taylor, W.J., 2017. Apps for people with rheumatoid arthritis to monitor their disease activity: a review of apps for best practice and quality. JMIR mHealth uHealth 5, e7. https://doi.org/10.2196/mhealth.6956.

Grundy, Q., Chiu, K., Held, F., Continella, A., Bero, L., Holz, R., 2019. Data sharing practices of medicines related apps and the mobile ecosystem: traffic, content, and network analysis. BMJ 364, l920. https://doi.org/10.1136/bmj.l920.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6, 65–70.

Huckvale, K., Torous, J., Larsen, M.E., 2019. Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. JAMA Netw. Open 2, e192542. https://doi.org/10.1001/jamanetworkopen.2019.2542.

Jackson, T., Thomas, S., Stabile, V., Shotwell, M., Han, X., McQueen, K., 2016. A systematic review and meta-analysis of the global burden of chronic pain without clear etiology in low- and middle-income countries: trends in heterogeneous data and a proposal for new assessment methods. Anesth. Analg. 123, 739–748. https://doi.org/10.1213/ANE.0000000000001389.

James, S.L., Abate, D., Abate, K.H., Abay, S.M., Abbafati, C., Abbasi, N., Abbastabar, H., et al., 2018. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 392, 1789–1858. https://doi.org/10.1016/S0140-6736(18)32279-7.

Knitza, J., Tascilar, K., Messner, E.-M., Meyer, M., Vossen, D., Pulla, A., Bosch, P., et al., 2019. German mobile apps in rheumatology: review and analysis using the mobile application rating scale (MARS). JMIR mHealth uHealth 7, e14991. https://doi.org/10.2196/14991.

Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J. Chiropr. Med. 15, 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Kumar, N., 2007. WHO normative guidelines on pain management report of a Delphi study to determine the need for guidelines that should be developed by WHO. World Heal. Organ. 1–50. doi:https://doi.org/10.1111/j.1526-4637.2012.01414.x.

Lalloo, C., Jibb, L.A., Rivera, J., Agarwal, A., Stinson, J.N., 2015. "There's a pain app for that": review of patient-targeted smartphone applications for pain management. Clin. J. Pain 31, 557–563. https://doi.org/10.1097/AJP.0000000000000171.

Lalloo, C., Shah, U., Birnie, K.A., Davies-Chalmers, C., Rivera, J., Stinson, J., Campbell, F., 2017. Commercially available smartphone apps to support postoperative pain self-management: scoping review. JMIR mHealth uHealth 5, e162. https://doi.org/10.2196/mhealth.8230.

Leadley, R.M., Armstrong, N., Lee, Y.C., Allen, A., Kleijnen, J., 2012. Chronic diseases in the European Union: the prevalence and health cost implications of chronic pain. J. Pain Palliat. Care Pharmacother. 26, 310–325. https://doi.org/10.3109/15360288.2012.736933.

Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., Fuller-Tyszkiewicz, M., 2019. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 18, 325–336. https://doi.org/10.1002/wps.20673.

Luxton, D.D., McCann, R.A., Bush, N.E., Mishkind, M.C., Reger, G.M., 2011. mHealth for mental health: integrating smartphone technology in behavioral healthcare. Prof. Psychol. Res. Pract. 42, 505–512. https://doi.org/10.1037/a0024485.

Machado, G.C., Pinheiro, M.B., Lee, H., Ahmed, O.H., Hendrick, P., Williams, C., Kamper, S.J., 2016. Smartphone apps for the self-management of low back pain: a systematic review. Best Pract. Res. Clin. Rheumatol. 30, 1098–1109. https://doi.org/10.1016/J.BERH.2017.04.002.

Mani, M., Kavanagh, D.J., Hides, L., Stoyanov, S.R., 2015. Review and evaluation of mindfulness-based iPhone apps. JMIR mHealth uHealth 3, e82. https://doi.org/10.2196/mhealth.4328.

Masterson Creber, R.M., Maurer, M.S., Reading, M., Hiraldo, G., Hickey, K.T., Iribarren, S., 2016. Review and analysis of existing mobile phone apps to support heart failure symptom monitoring and self-care management using the mobile application rating scale (MARS). JMIR mHealth uHealth 4, e74. https://doi.org/10.2196/mhealth.5882.

Messner, E.-M., Probst, T., O'Rourke, T., Baumeister, H., Stoyanov, S., 2019. mHealth applications: potentials, limitations, current quality and future directions. In: Baumeister, H., Montag, C. (Eds.), Mobile Sensing and Digital Phenotyping: New Developments in Psychoinformatics. Springer, Berlin.

Meßner, E., Terhorst, Y., Sander, L.B., Schultchen, D., Plaumann, K., Sturmbauer, S., Barke, A., et al., 2019. "When the fear kicks in"- Standardized Expert Quality Ratings of Apps That Aim to Reduce Anxiety.

Messner, E.-M., Terhorst, Y., Barke, A., Baumeister, H., Stoyanov, S., Hides, L., Kavanagh, D., et al., 2019b. Development and Validation of the German Version of the Mobile Application Rating Scale (MARS-G) (JMIR m u Heal. accepted).

Messner, E.-M., Terhorst, Y., Sander, L.B., Schultchen, D., Portenhauser, A., Schmidbaur, S., Stach, M., et al., 2021. A Systematic Review and Evaluation of Quality for Mobile Health Apps for the Management of Gastrointestinal Diseases (Prep).

Nicholas, J., Larsen, M.E., Proudfoot, J., Christensen, H., 2015. Mobile apps for bipolar disorder: a systematic review of features and content quality. J. Med. Internet Res. 17, e198 https://doi.org/10.2196/jmir.4581.

Paganini, S., Terhorst, Y., Sander, L.B., Catic, S., Balci, S., Küchler, A.-M., Schultchen, D., et al., 2019. "Move It!" Standardised Expert Quality Ratings (MARS) of Apps That Foster Physical Activity for Android and iOS (Prep).

Portenhauser, A., Terhorst, Y., Schultchen, D., Sander, L.B., Denkinger, M., Waldherr, N., Dallmeier, D., et al., 2021. A systematic review and evaluation of mobile applications for the elderly. JMIR Aging in press. https://doi.org/10.2196/23313.

Priebe, J.A., Haas, K.K., Moreno Sanchez, L.F., Schoefmann, K., Utpadel-Fischler, D.A., Stockert, P., Thoma, R., et al., 2020. <p>Digital treatment of back pain versus standard of care: the cluster-randomized controlled trial, rise-up</p>. J. Pain Res. Volume 13, 1823–1838. https://doi.org/10.2147/JPR.S260761.

Rathner, E.-M., Probst, T., 2018. Mobile Applikationen in der psychotherapeutischen Praxis: Chancen und Grenzen. PiD - Psychother. im Dialog 19, 51–55. https://doi.org/10.1055/a-0592-0496.

Rod, K., 2016. Finding ways to lift barriers to care for chronic pain patients: outcomes of using internet-based self-management activities to reduce pain and improve quality of life. Pain Res. Manag. 2016. https://doi.org/10.1155/2016/8714785.

Salazar, A., de Sola, H., Failde, I., Moral-Munoz, J.A., 2018. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the mobile app rating scale. JMIR mHealth uHealth 6, e10718. https://doi.org/10.2196/10718.

Sander, L.B., Schrondanner, J., Terhorst, Y., Spanhel, K., Pryss, R., Baumeister, H., Messner, E.-M., 2019. Help for trauma from the app stores? A systematic review and standardised rating of apps for post-traumatic stress disorder (PTSD). Eur. J. Psychotraumatol. accepted.

Sander, L.B., Schorndanner, J., Terhorst, Y., Spanhel, K., Pryss, R., Baumeister, H., Messner, E.-M., 2020. 'Help for trauma from the app stores?' A systematic review and standardised rating of apps for post-traumatic stress disorder (PTSD). Eur. J. Psychotraumatol. 11, 1701788. https://doi.org/10.1080/20008198.2019.1701788.

Santo, K., Richtering, S.S., Chalmers, J., Thiagalingam, A., Chow, C.K., Redfern, J., 2016. Mobile phone apps to improve medication adherence: a systematic stepwise process to identify high-quality apps. JMIR mHealth uHealth 4, e132. https://doi.org/10.2196/mhealth.6742.

Schoeppe, S., Alley, S., Rebar, A.L., Hayman, M., Bray, N.A., Van Lippevelde, W., Gnam, J.-P., et al., 2017. Apps to improve diet, physical activity and sedentary behaviour in children and adolescents: a review of quality, features and behaviour change techniques. Int. J. Behav. Nutr. Phys. Act. 14, 83. https://doi.org/10.1186/s12966-017-0538-3.

Schultchen, D., Terhorst, Y., Holderied, T., Stach, M., Messner, E.-M., Baumeister, H., Sander, L.B., 2020. Stay present with your phone: a systematic review and standardized rating of mindfulness apps in European app stores. Int. J. Behav. Med. 1–9.

Shen, N., Levitan, M.-J., Johnson, A., Bender, J.L., Hamilton-Page, M., Jadad, A. (Alex) R., Wiljer, D., 2015. Finding a depression app: a review and content analysis of the depression app marketplace. JMIR mHealth uHealth 3, e16. https://doi.org/10.2196/mhealth.3713.

StatCounter, 2019. Mobile Operating System Market Share Worldwide [WWW Document]. URL. https://gs.statcounter.com/os-market-share/mobile/worldwide.

Stoyanov, S.R., Hides, L., Kavanagh, D.J., Zelenko, O., Tjondronegoro, D., Mani, M., 2015. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR mHealth uHealth 3, e27. https://doi.org/10.2196/mhealth.3422.

Stoyanov, S.R., Hides, L., Kavanagh, D.J., Wilson, H., 2016. Development and validation of the user version of the mobile application rating scale (uMARS). JMIR mHealth uHealth 4, e72. https://doi.org/10.2196/mhealth.5849.

Sucala, M., Cuijpers, P., Muench, F., Cardoș, R., Soflau, R., Dobrean, A., Achimas-Cadariu, P., et al., 2017. Anxiety: there is an app for that. A systematic review of anxiety apps. Depress. Anxiety 34, 518–525. https://doi.org/10.1002/da.22654.

Terhorst, Y., Rathner, E.-M., Baumeister, H., Sander, L.B., 2018. "Help from the app store?": a systematic review of depression apps in the German app stores. Verhaltenstherapie 28.

Terhorst, Y., Philippi, P., Sander, L.B., Schultchen, D., Paganini, S., Bardus, M., Santo, K., et al., 2020. Validation of the mobile application rating scale (MARS). PLoS One 15, e0241480. https://doi.org/10.1371/journal.pone.0241480.

THE EUROPEAN PARILIAMENT, THE COUNCIL OF THE EUROPEAN UNION, 2018. Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (general data protection regulation – GDPR). Off. J. Eur. Union 2014, 958–981. https://doi.org/10.5771/9783845266190-974.

Thornton, L., Quinn, C., Birrell, L., Guillaumier, A., Shaw, B., Forbes, E., Deady, M., et al., 2017. Free smoking cessation mobile apps available in Australia: a quality review and content analysis. Aust. N. Z. J. Public Health 41, 625–630. https://doi.org/10.1111/1753-6405.12688.

Thurnheer, S.E., Gravestock, I., Pichierri, G., Steurer, J., Burgstaller, J.M., 2018. Benefits of mobile apps in pain management: systematic review. JMIR mHealth uHealth 6, e11231. https://doi.org/10.2196/11231.

Toelle, T.R., Utpadel-Fischler, D.A., Haas, K.-K., Priebe, J.A., 2019. App-based multidisciplinary back pain treatment versus combined physiotherapy plus online education: a randomized controlled trial. npj Digit. Med. 2, 34. https://doi.org/10.1038/s41746-019-0109-x.

Torous, J., Andersson, G., Bertagnoli, A., Christensen, H., Cuijpers, P., Firth, J., Haim, A., et al., 2019. Towards a consensus around standards for smartphone apps and digital mental health. World Psychiatry 18, 97–98. https://doi.org/10.1002/wps.20592.

Traue, H., Horn, A.B., Kessler, H., Jerg-Bretzke, L., 2013. Psyche und Schmerz, in: Baron, R., Koppert, W., Strumpf, M., Willweber-Strumpf, A. (Eds.), Praktische Schmerzmedizin. Interdisziplinäre Diagnostik - Multimodale Therapie. pp. 117–128.

Turk, D.C., Wilson, H.D., Cahana, A., 2011. Treatment of chronic non-cancer pain. Lancet 377, 2226–2235. https://doi.org/10.1016/S0140-6736(11)60402-9.

Veehof, M.M., Trompetter, H.R., Bohlmeijer, E.T., Schreurs, K.M.G., 2016. Acceptance- and mindfulness-based interventions for the treatment of chronic pain: a meta-analytic review. Cogn. Behav. Ther. 45, 5–31. https://doi.org/10.1080/16506073.2015.1098724.

Wallace, L.S., Dhingra, L.K., 2014. A systematic review of smartphone applications for chronic pain available for download in the United States. J. Opioid Manag. 10, 63. https://doi.org/10.5055/jom.2014.0193.

Weisel, K.K., Fuhrmann, L.M., Berking, M., Baumeister, H., Cuijpers, P., Ebert, D.D., 2019. Standalone smartphone apps for mental health—a systematic review and meta-analysis. npj Digit. Med. 2, 118. https://doi.org/10.1038/s41746-019-0188-8.