Available mHealth Applications for Spinal Surgery Patients: A Systematic App Review.

by

Jennifer Schulze (s1980955)

A Thesis Submitted in Partial Fulfilment of the Requirements for the Degree of

Bachelor of Science

Supervisors:

Annemieke Y. van der Horst MSc.

Dr. Saskia M. Kelders

THE UNIVERSITY OF TWENTE

Faculty of Behavioural, Management and Social Sciences

Department of Psychology, Health, and Technology

July, 2020

Abstract

Despite the broad availability of mHealth applications in the spinal surgery sector, their content, quality, and scientific evidence remains vague. Therefore, a systematic review was conducted to assess mHealth applications targeted to spinal surgery patients and spinal surgery interventions. The two major app stores (IOS and Android) were searched. Apps were evaluated using the Mobile App Rating Scale (MARS) along with a qualitative content assessment. In addition, available scientific evidence regarding the apps' effectiveness was collected. 43 unique apps were identified focussing on physical health, well-being, exercise, and education. On a 5-point scale, the overall app quality mean score was 3.64, with individual app scores ranging from 4.76 to 2.05. The overall in-app engagement was poor, with a total mean score of 3.16. Apps scored highest in the category 'Functionality' with a total mean score of 4.16. In addition, it was found that the scientific evidence of the assessed apps was limited, with only 5 apps supported by scientific evidence. Therefore, this review highlights the need for more scientific testing regarding mHealth apps' effectiveness, along with quality improvement regarding user engagement.

Table of Contents

Introduction……………………………………………………………………………………….1

Methods…………………………………………………………………………………………...4

Study Design…………………………………………………………………………….....4

Search for available apps………………………………………………………………...4

Inclusion and exclusion criteria……………………………………………………….....6

Data extraction and analysis……………………………………………………………..6

MARS evaluation………………………………………………………………………..8

Scientific evidence and effectiveness…………………………………………………....8

Results……………………………………………………………………………………….9

App store search…………………………………………………………………………9

Descriptive features of the apps…………………………………………………………10

App content analysis…………………………………………………………………… 11

Outcome measures of the Mobile app rating scale……………………………………….. 14

App classification………………………………………………………………………. 14

App quality ratings………………………………………………………………………15

Scientific evidence and effectiveness……………………………………………………18

Mentionable apps……………………………………………………………………….. 20

Discussion…………………………………………………………………………………21

Interpretation……………………………………………………………………………..22

Implications………………………………………………………………………………23

Limitations………………………………………………………………………………..23

Recommendations……………………………………………………………………….. 24

Conclusion………………………………………………………………………………….25

References………………………………………………………………………………… 26

Appendix……………………………………………………………………………………...31

Appendix A: Collected Features and in-app Functions…………………………………….31

Appendix B: Scientific Evidence of mHealth Apps………………………………………..35

Appendix C: Examples of well-designed Applications…………………………………….38

Appendix D: Examples of bad designed Applications……………………………………..45

Appendix E: Categorical App Examples…………………………………………………...47

Appendix F: App Duplicates………………………………………………………………..58

Introduction

In today's day and age, the constant use of smartphones established itself as a vital part of our every day- and communal life. Along with the technological advancements, more and more aspects of our lives are shifting into the digital world, including health and the way healthcare is delivered. This intention is closely related to the concept of eHealth which refers to the 'use of technology to improve health, well-being and healthcare' and the 'use of information and communication technologies (ICT) for health' (van Gemert-Pijnen, Kelders, Kip, & Sanderman, 2018; WHO, 2019).

As a corresponding part of eHealth, mHealth (or mobile health) describes the use of smart or portable devices for health services and information, usually mediated through technologies such as smartphones, tablets, smartwatches, and other wearables (Moss, Süle, & Kohl, 2019). With the mission to improve the way healthcare is delivered, mHealth targets the enhancement of patient self-care and management, along with the goal of improving patient's quality of life (Malvey & Slovensky, 2014; Oh et al., 2015). Facilitated through technological development, enhanced connectivity, and availability and accessibility, mHealth creates a new way of convenient healthcare benefiting both healthcare professional and patients. In addition, mHealth holds the potential to a significant societal impact, including not only enhanced patient well-being, but also quick facilitation of patient information, the inclusion of undersupplied population groups, and significant resource and cost reduction. (Hou et al.,2019; van Gemert-Pijnen et al., 2018; Waller et al., 2015).

Contingent upon the advantages of mHealth, many healthcare sectors initiated the development of many mHealth-related apps to improve the way healthcare is delivered. Therefore, it is unsurprising that in the year 2017, around 325.000 mHealth apps were available at the two major app stores (Pohl, 2017). One of the sectors that is especially profiting from development of mHealth application is the sector of spinal surgery, where mHealth is already viewed as a new high-tech solution for enhancing spinal surgery treatment (Semple, Sharpe, Murnagn, Theodoropoulos, & Metcalfe, 2015). With spinal surgery interventions exceeding 1.94 procedures and expenses of $200 billion dollars every year in the US, healthcare professionals and the sector get increasingly overallocated, failing to meet the needs of the patient and ensure a complication-free and smooth surgical intervention (Vainshnav & McAnany, 2019; Waller et al., 2015). This does not only result in cancellation, delays, medical problems readmissions, and longer hospital stays, but also in an opportunity for the implementation of mHealth, tackling these issues (Andersson & Watkins-Castillo, 2014; Waller et al., 2015).

Consequently, mHealth applications targeting spinal surgery patients started to emerge, along with studies testing their viability. In a study of Debono and colleagues (2016), the feasibility of a mHealth application for postoperative monitoring targeted to outpatient lumbar discectomy patients was assessed. The app featured digital follow-up monitoring, where the patient received daily questionnaires and checklists about his physical condition, monitored by nurses. The study resulted in a reduction of hospital time, better patient monitoring, care management and overall effectiveness of mHealth monitoring in chronic disease management (Debono et al., 2016).

In another study, Semple and his colleagues (2015) reported that postoperative surgery complications could be identified by using patients' photos of surgical wounds mediated through the app. Providing surgeons with the opportunity of monitoring the patients more effectively, complications could be earlier detected, as well as the prevention and reduction of unexpected readmissions and follow-up appointments (Semple et al., 2015). Along with that, Waller and colleagues (2015) examined preoperative eHealth education programs about orthopaedic surgery interventions, including spinal surgery. It was found that these programs can reduce patient anxiety, enhance patient compliance and adherence, and satisfaction in both the patients and surgeons (Waller et al., 2015).

Although the studies regarding mHealth apps targeting spinal surgery might be promising, the implementation into the treatment process and surgical interventions in hospitals is still missing. With mHealth still pioneering, there are no current existing rules regarding content, medical accuracy, and effectiveness of apps (Dawson et al., 2019). Since the opportunity to develop and publish healthcare is given to anyone, it gets increasingly confusing to detect apps exhibiting medical involvement and suitability for medical application (Rajak & Shaw, 2019). This is rising issue, that complicates not only the search for available evidence of the quality and quality of spinal surgery apps but also decreases the number of trustworthy and high-quality apps available (Dittrich et al., 2020). Consequently, mHealth applications containing outdated practice guidelines, technological features, and misleading content are dominating. This not only places the user at risk but also impedes the implementation of mHealth applications into clinical practice (Dawson et al., 2019; Malvey & Slovensky, 2014).

Because there are no rules that guide the development of those apps, assessment of app content and quality lies within health researchers. As a result, app store screenings and app-ratings, like the MARS, emerged testing the content and quality of mHealth apps (Jake-Schoffmann et al., 2017; Stoyanov et al., 2015). However, systematic reviews as well as research regarding content, quality, and evidence of mHealth apps are still widely unavailable in all sectors of surgery- with no current review conducted on mHealth apps addressing spinal surgery (Dawson et al., 2019; Dittrich et al., 2020; Semple et al., 2015).

The resulting lack of systematic app review in the sector of spinal surgery creates a barrier for successful implementation of mHealth into the treatment process (Istepanian & Woodward, 2016). To enhance the way healthcare is delivered in spinal surgery and to clear the way for successful implementation in clinical practice, a content evaluation is needed to assess whether the current mHealth apps provide useful content information and are safe in clinical practice (Dittrich et al., 2020).

In order to remove this barrier and raise the awareness of existing mHealth interventions in the sector of spinal surgery, a systematic app store review in the sector of spinal surgery needs to be conducted to find available mHealth interventions, while examining their content and quality, and scientific evidence. To solve this problem, this systematic review addresses the question: ‘What is the content, quality, and evidence of spinal surgery interventions present in currently available mHealth apps targeted at spinal surgery patients?’

Methods

Study design

To answer this question, a systematic mHealth app review was performed. The two major app stores, IOS and Android, were searched for apps aimed at spinal surgery patients and their surgical intervention. Afterwards, a quality assessment was made addressing the app content, quality, and scientific background. For the content evaluation, a qualitative method was used to evaluate features, functions, and characteristics of available applications. App quality was assessed using the Mobile App Rating Scale (MARS)- a reliable, expert-based, multidimensional scale designed for classifying and rating the quality of mobile health apps (Stoyanov et al., 2015). The scientific background of apps as well as their effectiveness was assessed by searching and screening app developer’s websites and databases.

Search for available apps

Search and Risk of Bias. The search was conducted in a one-month period from April 2020- May 2020. The IOS and Android stores were searched using an iPhone 6s (Software version 13.3.1), a Huawei P20 Pro (Software version EMUI 9.1.0.380) and an iPad Pro 11 (Software version 13.4). To avoid any bias or exclusion due to app store algorithms and personalized offers, new accounts on all devices were created. Moreover, to prevent exclusion of region- and language-specific apps, a VPN connection was used to conduct the search from

Germany, the UK, the Netherlands, USA, Australia, and Canada. In addition, the storage of all devices was cleared to ensure a smooth run of the apps and prevent complications within the content and quality assessment.

Search Procedure. In the first round of search, apps in both stores were searched using keywords related to the research question, such as '*spinal surgery*' or '*back surgery recovery*'. In a second round, terms generated from app descriptions and titles were used in both app stores to refine the search regarding single spinal conditions. For example, '*herniated disc*' used for degenerative spine diseases, '*spinal vertebrae fractures*' for spinal trauma, and '*scoliosis*' for spinal deformities. If a suitable app was found, snowballing was used to find other apps suggested on the related page. The overview of the generated search terms and results can be found in Table 1.

Table 1

Results of the App Store search terms

	App Store		Google Play Store	
Search Terms	N	relevant	N	relevant
Surgery recovery	26	18	29	17
Spinal Surgery	106	7	144	38
Spine surgery	86	12	216	54
Back surgery	107	9	123	36
Back surgery recovery	0	0	198	46
Back pain	180	82	238	44
Rug operatie revalidatie	0	0	147	42
Rug operatie	1	1	221	45
Rücken operation	2	1	224	59
rückenoperation	1	1	106	54
Rücken-op	85	15	-	-
Spinal surgery recovery	0	0	250	32
Rücken reha	3	2	204	24
Wirbelsäulen operation	1	1	142	51

	App Store		Google Play Store	
Search Terms	N	relevant	N	relevant
Wirbelsäule reha	0	0	134	37
Herniated disc	0	0	124	45
Lumbar spinal stenosis	0	0	89	34
Spinal vertebrae fractures	0	0	73	40
scoliosis	30	18	118	52
malformation	4	0	79	0
Spinal deformity	6	4	121	56
Spinal tumor	2	1	-	-
Bandscheibenvorfall	3	2	-	-

Note: the numbers do not display the unique numbers of apps found.

Inclusion and Exclusion Criteria

The following inclusion criteria needed to be met: (a) the apps featured the languages English, Dutch, and German; multi-language apps were included when one of the target languages was manually adjustable in the app settings. (b) Apps free of charge, (c) accessible apps; if an app was downloadable but not accessible in terms of an internal error, a missing license, password, code, or other means of authorization, it was excluded. (d) Apps related to spinal surgery patients and spinal surgery interventions in terms of secondary prevention (pre- and postoperative care). (e) Apps solely made for wearable devices were excluded because wearables were unavailable for the review. (f) Apps containing games, wallpaper apps, and journals were excluded. (g) Apps updated longer than five years ago were excluded, due to uncertainty whether the app's content reflects the current medical standards (Dawson et al., 2019).

Data extraction and analysis

For the data extraction, the following steps were conducted: (1) the apps were searched and screened on their title and picture. If the app seemed to be suitable, preliminary screening of potential apps for relevance and inclusion was conducted based on the app description, screenshots, reviews, update, language, and device comparability. (2) The app was downloaded if it (for the time and being) matched the inclusion criteria and subsequently checked on accessibility. When the application was not accessible, an email was sent to the developer requesting access. (3) If the downloaded app was accessible, it was again checked on suitability and excluded if it did not match the inclusion criteria. (4) If the app was accessible and matched the inclusion criteria, a short cognitive walkthrough was performed assessing app focus (e.g. if the app focussed on the pre-operative phase, postoperative phase, or whole intervention), the concept of the app (e.g. if the app targeted exercise, information, surgery preparation), features, and in-built functions.

Evaluation of Data. The collected information was entered into a table and summarized into four categories: application name and developer, descriptive data, app content, and additional annotations (e.g. if the app needed access, if the app was provided by a particular clinic or individual healthcare professional). During this process, eight duplicates were detected, all developed by the same company. Since the information gathered as well as content for these apps was identical (except colour, clinic-specific logo, number of reviews and downloads), the ‘original’ app from the developer, CASPAR-health, was included in the search with the duplicates being excluded from the evaluation. A summary of the duplicates can be found in Appendix F.

Descriptive features and app content analysis. Once the search was completed, the gathered information was separated into two different sections. The first section summarized the descriptive features of the app, including app category and focus (e.g. if the app focussed on the pre- or postoperative phase, or the whole intervention), advertisement and updates (whether the app contained any form of advertisement or in-app purchases), access and device compatibility (whether the app was related to a specific hospital or healthcare professional; compatibility to external devices, e.g. wearables), and language.

The second section addressed the content of the apps. The previously gathered information was evaluated by highlighting similar content among the apps. Inspired by the design suggestions in the research of Mendiola and colleagues (2015) and Bendixen and colleagues (2017), ten different categories were created, determined by the most common characteristics, features, and in-built functions among the apps. The categories are further explained with the results of the app content analysis.

MARS evaluation

To assess the app quality, the Mobile App Rating Scale (MARS) was used. It comprises three sections: classification, quality, and satisfaction. The classification section provides descriptive information about the app (e.g. rating, update, version, platform, app target (focus), theoretical background/strategies (e.g. whether the app used assessment, monitoring, tracking, goal setting, was strength-based), affiliations, age group, technical aspects). The app quality section includes 19 items, assessing the quality on four dimensions: User Engagement (Section A), App functionality (Section B), Aesthetics (Section C), and Information quantity and quality (Section D). MARS items are scored using a 5-point Likert scale, from 1- ‘inadequate’ to 5 ‘excellent’. The third section, satisfaction (Section E), assessed the subjective rating of the app.

The assessment was made using the downloaded apps, as well as the already collected information. The mean scores of the four subscales (Section A-D) were calculated. However, scientific evidence (Item 19, Section D), was excluded because it was separately addressed and in a majority of cases non-applicable. The final app quality mean scores were calculated, along with the subjective quality score. Lastly, results were ranked on their app quality mean score. For the app classification section, frequencies were calculated.

Scientific evidence and app effectiveness

After the evaluation, the scientific evidence regarding the effectiveness of the assessed apps was addressed. The developers of all apps were contacted, using the provided information on the website of the specific app store. An email was sent regarding the information addressed in item 19 of the MARS (Section D). It was asked whether the developed app underwent usability testing or was tested regarding its effectiveness and validity in published scientific literature. Additionally, the app description, as well as the developer's website, was screened for any scientific evidence and publications of the app. In case the developer left the email unanswered, several databases (Google Scholar, PubMed, Science Direct, LISA, Medline, Cochrane Library) were searched for scientific publications instead. The effectiveness of apps was assessed by screening the articles and extracting information.

Results

App Store Search

The app store search resulted in a total of 842 records: 563 in the Google Play Store and 279 in the App store (Figure 1). After excluding the liable to pay applications, apps that were not on topic (e.g. games) and duplicates among both app stores, 213 unique records were screened on its title and description. A total of 132 applications were sorted out as they were not on topic, had a different language, were incompatible with the available devices, or were updated longer than 5 years ago. A total of 81 apps were downloaded, from which 31 apps were removed after downloading due to access issues, being a duplicate (CASPAR) or containing in-app issues, for example, inaccessible app components which did not load, or an app crash, resulting in app assessment complications.

Figure 1. Flow chart depicting the selection of procedure of applications.

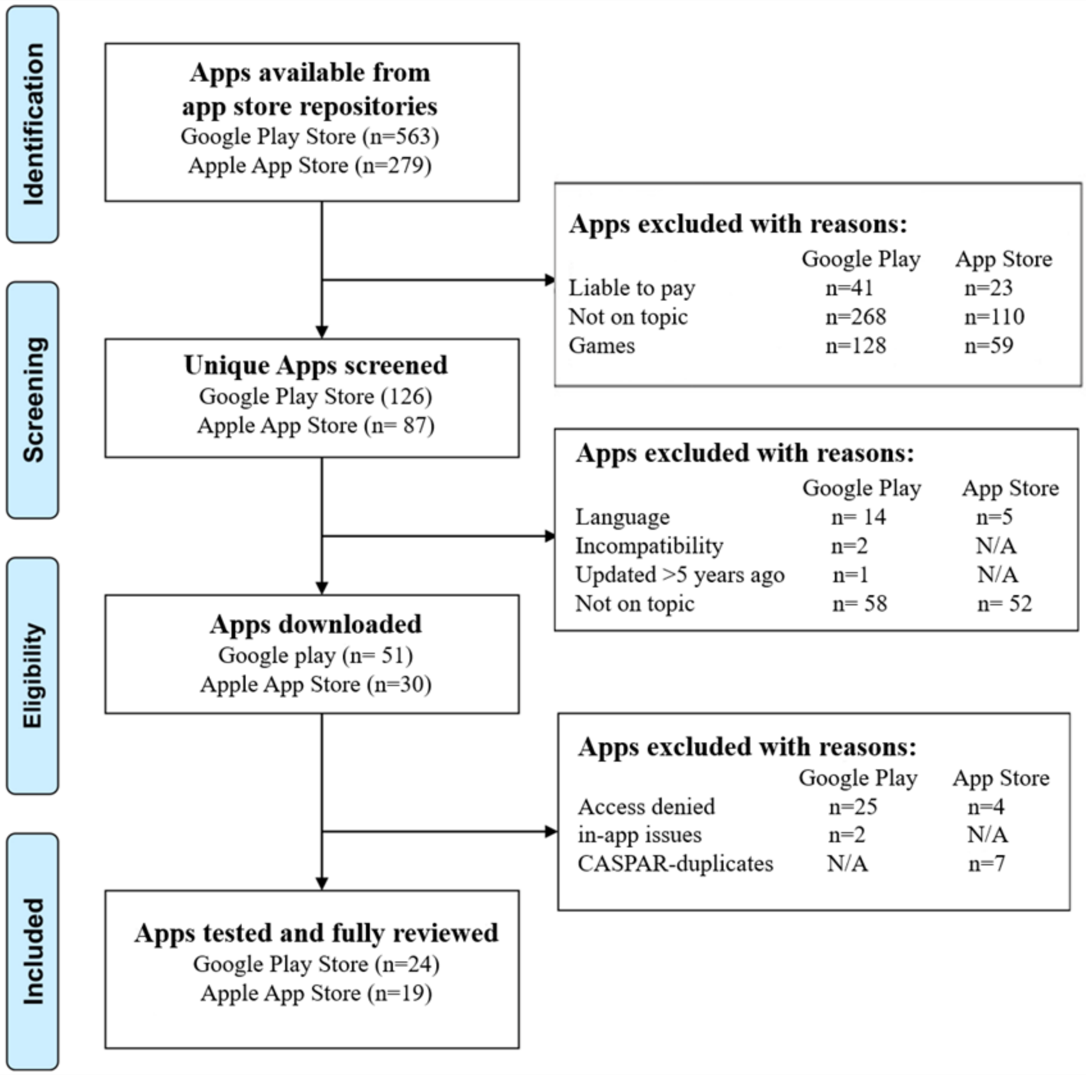

Figure 1. Process of searching, extracting, and selecting apps from the Google Play Store and Apple App Store.

Descriptive features of the apps

For the final assessment, 43 unique applications from both app stores were included. 24 Apps (55.81%) were found on the Google Play Store, while 19 apps (44.18%) were found on the Apple Store. 27 apps (62.79%) were present in both stores.

App Categories and Focus. The apps can be sorted into 2 categories based on the app-store classification, namely ‘medical’, related to 21 Apps (48.83%) and ‘Health & Fitness’, related to 22 apps (51.16%). Regarding the applications’ target and focus, 29 apps focused on the postoperative phase (67.44%), while 9 apps focused on the preoperative phase (20.93%). 5 apps focused enclosed the whole surgical intervention (11.62%).

Advertisement and Updates. Although all apps were free of charge, 6 apps (13.95%) involved minor in-app purchasing, for example additional workout plans. Furthermore, another 6 apps (13.95%) included advertisement, which was depicted as a banner at the top or the bottom of the app screen. Most apps were up to date, with 12 apps (27.90%) updated less than 1 month ago, 3 apps (6.97%) updated in 2020, 17 apps (39.53%) in 2019, 3 apps (6.97%) in 2018, 6 apps (13.95%) in earlier years. 2 apps were still in its first version (4.65%).

Access and Device Compatibility. 17 applications required access (39.53%) with 17 apps (39.53%) related to a specific clinic, hospital, or therapy of a healthcare professional. Among the 43 apps, only 4 apps could be synched to a specific device (9.3%), namely wearables via apple health kit (2 apps), an ABS- device (1 app), and a device which can be attached to the phone to measure spinal deformity (1 app).

App languages. Most of the apps were only available in the English language (58.19%). In addition, 4 apps were only available in German (9.3%). Out of all apps, 14 apps featured multiple languages besides English and German, namely French (6), Dutch (3), Spanish (3), Arabic (3), Chinese (2), Danish (2), Finnish (2), Swedish (2), Norwegian (2), and Russian (2).

App Content Analysis

The 10 categories formed are labelled ‘SNS’, ‘Multipurpose’, ‘Rewards’, ‘Medication’, ‘Reminders’, ‘Pain Assessment’, ‘Datasharing’, ‘Progress’, ‘Checklist’, and ‘Motivation’. The definitions, along with the number of apps matching the different categories, are summarized in Table 2.

SNS and Data sharing. 9 Apps contained a messenger system, where patients could message their healthcare providers, ask them questions, or send them information related to their surgical intervention. 1 app involved a social network circle, where patients could connect themselves to ask questions, share their progress and pain (Appendix E, Figure 21 ‘CatchMyPain Community’). Along with the opportunity to message healthcare providers via an in-built SNS, 18 apps provided the user with the possibility to share app data with healthcare professionals or stakeholders on other platforms. For example, the app data and records could be exported via pdf or email or could be shared on other social media networks if desired.

Multipurpose. 16 apps targeted not only spinal surgery patients and the spinal surgery intervention, but also included patients from other sectors, for example knee, abdominal, shoulder, and hip surgery patients. This was especially present in apps focussing on recovery, with several exercises suitable for different patients.

Rewards and Motivation. 3 apps also provided the users with achievements or points for a completed task. Rewards were given after a completed exercise unit, or another accomplished task mostly related to surgery recovery. Along with giving rewards, 5 apps provided a motivational component to engage the user. This was usually present after a completed exercise, questionnaire, or task (e.g. ‘you did it’, ‘keep going’, ‘you are almost there’) (Appendix C, Figure 6 ‘Mineo’; Appendix C, Figure 7 ‘Kaia’).

Medication. 4 apps incorporated a medication feature, enabling the patient to enter their medications, doses, name of the medication, and the time for taking medications. In 2 clinic-related apps, this information was entered by healthcare professionals working at the hospital. An example can be found in Appendix E, Figure 23 ‘Postapp’.

Reminders. 15 Apps incorporated reminders presented as push-notifications on the phone. In all apps, this function could be adjusted by the user. The reminders targeted medication intake, completion of pain assessments, questionnaires, or exercise units. Furthermore, reminders could also be set to engage the user in completing essential steps regarding the intervention process. (Appendix C, Figure 5. Vivira ‘Stay on Track with exercise reminders’)

Pain Assessment and Progress Monitoring. 19 apps featured method of pain assessment, in forms of scales, diary or pain assessment tools. For example, users could assess their general pain or back pain on a 10-point scale, labelled by emoticons and descriptions of the pain. Some also featured a diary where the user could write something about his progress, pain or feelings. (Appendix E, Figure 15; Appendix E, Figure 20 ‘myrecovery’). Along with the opportunity to pain assessment, 22 apps enabled the patient to monitor his progress. Most common were statics about pain, completed exercise/ exercise progress, mental well-being, and an exercise history (Appendix E, Figure 20 ‘myrecovery’).

Checklists. 21 Apps provided checklists or questionnaires regarding the surgical intervention, spinal condition, the hospital, or the patient oneself. For example, some clinic-related applications provided checklists on what to bring or what do prepare before a surgical appointment, guiding the patient through all the steps he has to fulfil before the surgical intervention, or during recovery.

Table 2.

App Categories and Definitions

Category Name	Definition	N Apps
SNS	The app contains an in-built social network circle or a messenger function	10
Multipurpose	Besides targeting spinal surgery patients, the app includes patients from other orthopaedic sections (e.g. knee surgery) as well	16
Rewards	The app provides users with rewards (e.g. achievements, points)	3
Medication	The app provides the opportunity of monitoring or entering prescribed medication	4
Reminders	The app sends the user reminders (e.g. push- notifications)	15
Pain Assessment	The app provides quantitative or qualitative methods (scales, diary) to assess the pain of the user	19
Data-sharing	The app provides the possibility to share the app data with doctors or other stakeholders (e.g. via pdf, email)	18

Category Name	Definition	N Apps
Progress	The app provides a method of monitoring progress or pain	22
Checklist	The app provides checklists or questionnaires regarding the surgical intervention, spinal condition, hospital, or the patient himself	21
Motivation	The app involves a motivational component to engage the user (e.g. ‘you did it’ after an exercise component)	5

Outcome Measures of the Mobile App Rating Scale (MARS)

App Classification

In total, 27 apps (62.79%) had app-user reviews, while 16 apps (37.20%) had no reviews. With an average customer rating of n= 211, 20 apps (46.51%) were below 100 ratings, 6 apps (13.95%) were rated by 100 customers or more, and 3 apps (6.97%) were rated by more than 1000 customers. The highest rating was n=32,657 at the point of evaluation.

Focus and App Target. Almost all assessed applications target physical health, , (n=42) concentrating on proper surgery recovery, the absence of disease, or increasing or maintaining the users’ fitness level. This category was followed by ‘increasing well-being’ (n=18), and other targets, namely ‘exercise’ (n=21) and ‘education’ (n=16). Some apps also focused on ‘goal setting’ (n=4), ‘reducing negative emotions’ (n=3) and ‘anxiety and stress’ (n=3).

Theoretical Background and Strategies. The theoretical background used in apps were mainly ‘information and education’ (n=25), followed by ‘exercise’ (n=21), including ‘strength-based’ exercises (n=14). Further theoretical background were ‘assessment’ (n=22), ‘monitoring/tracking’ (n=17), ‘advice/tips/strategies/skills training’ (n=16), ‘other’ (n=11), and ‘goal setting’ (n=5), focussing on setting and maintaining therapy goals.

Affiliations. 46.51% of the apps were affiliated to a small NGO or institution like a hospital or medical centre. 23.25% were affiliated to a commercial business, in another 23.25%

of the apps, the source could not be identified. 4.65% of the apps were associated with a University.

Age group. 51.16% of the Apps found in the App Stores were classified as PEGI 3, expressing suitability for the general population. 39.53% of the apps were suitable for customers aged 18 or older. 9.3% of the apps declared a usage for customers aged 13 or older.

Technical Aspects. Regarding the technical aspects of the app, the most common feature was the necessity for web access to function (n=24), followed by a required login (n=21) and an appertaining password protection (n=20). Another common technical aspect was the reminder function (n=24). A sharing option to social networks (n=4) and an in-app community (n=3) was relatively rare.

App Quality Ratings

The quality mean scores from the MARS evaluation ranged from 4,76 to 2,05. With an overall app quality $\bar{x}$=3.64, 35 apps (81.39%) reached the minimum acceptability score of 3.0. Out of four main sections of the MARS, apps scored lowest in Engagement ($\bar{x}$=3,17) and Information ($\bar{x}$=3.55). The apps scored highest in the category Functionality ($\bar{x}$=4.16) and Aesthetics ($\bar{x}$=3.64). The mean score for Subjective Quality was $\bar{x}$=3.64. Results of the 4 subscale scores along with the app quality mean and subjective quality are listed in Table 3.

Table 3

Mean Mobile App Rating Scale (MARS) Scores of the assessed Apps

App Name	Section A	Section B	Section C	Section D	App Quality Mean	Subjective Quality
Vivira: Back, Knee, and Hip	4.8	4.75	5.0	4.5	4.76	5.0
Kaia	4.8	4.75	5.0	4.3	4.71	4.75
mineo	4.4	5.0	5.0	4.3	4.67	5.0
myHealth Track	4.8	4.75	5.0	3.8	4.58	4.75

App Name	Section A	Section B	Section C	Section D	App Quality Mean	Subjective Quality
Doado, Your Back Companion	3.8	4.75	4.3	4.4	4.35	4.5
Allina Health	4.2	4.5	4.6	4.0	4.32	3.75
Ready Surgery App	4.0	4.5	4.6	4.2	4.32	3.75
CASPAR Health-Apps	4.0	4.5	4.3	4.16	4.24	4.5
Bauerfeind Therapie App	3.6	4.75	4.6	4.0	4.23	3.75
my recovery	3.0	4.75	4.6	4.3	4.17	3.5
Rally Recover	4.0	4.5	4.6	3.6	4.17	3.5
Spine Score	3.4	5.0	4.3	4.0	4.17	3.25
Risk Assessment Tool for Spine Surgery	3.4	4.75	4.0	4.5	4.16	3.5
BACS	4.0	4.75	3.6	4.0	4.08	3.25
Median- App	3.8	4.25	4.3	4.0	4.08	3.25
CatchMyPain	4.0	4.5	4.0	3.8	4.07	3.75
SpineScan3DR	3.2	4.75	4.0	4.16	4.02	4.5
INSELhealth-spinal surgery	3.6	3.75	4.0	4.4	3.93	4.5
MyChart	4.0	3.75	4.0	3.83	3.93	3.75
Kur-Reha BR	3.6	4.25	3.6	4.0	3.86	3.75
SORT	3.2	4.5	3.6	4.0	3.82	3.25
Pocket Spine Doc	3.4	4.25	4.0	3.6	3.81	3.5
iLog	3.8	4.0	4.0	3.6	3.8	4.0
Suxeed	2.8	4.75	4.3	3.16	3.75	3.25
THE SPINE APP	3.2	3.75	4.0	4.0	3.73	4.5
Motion Spine Institute	3.4	4.5	3.6	3.0	3.62	2.5
Back pain exercises	2.6	4.75	4.0	3.0	3.58	3.25
Rehab My Patient	3.0	4.25	2.5	4.0	3.43	4.25

App Name	Section A	Section B	Section C	Section D	App Quality Mean	Subjective Quality
Deuk Spine Institute	3.4	4.0	2.3	4.0	3.42	3.25
SLIC	2.4	4.5	3.0	3.5	3.42	2.5
Neurochirurgie	2.6	4.0	3.3	3.6	3.37	2.5
PostMed Patient App	2.6	4.25	3.6	3.0	3.36	3.0
Home Physio	2.4	3.75	3.3	3.4	3.21	3.0
Back Doctor (Free)	2.2	3.75	3.3	3.6	3.21	3.0
Lumbar herniated disc exercises	2.2	4.25	3.0	2.6	3.01	2.75
Hermiated Disc exercises	1.4	4.25	2.3	2.6	2.63	1.5
Spine & Tests	2.0	2.5	2.3	2.4	2.61	1.5
Spine health- spine, knee, and back pain workout	2.2	3.25	2.0	2.4	2.46	1.75
Lower back pain relief exercises	1.6	3.25	2.3	2.4	2.38	1.5
ShimSpine Exercises	2,6	2.25	2.3	2.0	2.33	1.25
Post Operative Solutions	2,4	2.2	2.5	2.0	2.27	2.16
herniated disc exercises	1.2	3.25	2.3	2.2	2.23	1.5
Sciatic Nerve Pain Exercises	1.2	3.5	1.3	2.2	2.05	1.0

Note: Caspar Health- APPS represents all 7 duplicates since they generated the same values.

Engagement. The mean scores for engagement ranged from $\bar{x}$=1.02 (Sciatic nerve pain exercises) to $\bar{x}$=4.8 (e.g. Vivira). Of the 43 apps, 23 scored high in matching its content with the target audience, presenting well-targeted visual information, design, and language for spinal surgery patients. However, many apps displayed insufficient interactivity, feedback, or lacked other input options with limited functions. Moreover, there were not enough strategies to increase engagement through entertainment.

Functionality. With mean scores ranging from $\bar{x}$=2.2 (Post-Operative Solutions) to $\bar{x}$=5 (e.g. Spine Score), the majority of apps displayed a high level of performance, including proper functionality regarding app features and components, e.g. buttons or the menu. Furthermore, the response time in many apps was fast, with no technical bugs. Moreover, the gestural design, as well as the navigation, was good. The apps were usable after some time and effort, displaying consistency between the single in-app screens.

Aesthetics. The means scores for aesthetics ranged from $\bar{x}=1.3$ (Sciatic Nerve Pain exercises) to $\bar{x}=5$ (e.g. Mineo). The graphic design, overall visual appeal, colour scheme and stylistic consistency was mostly clear and consistent in most apps, with no issues to select and locate single components or items of the apps. However, the picture as well as video quality differed in some apps, with some pictures being unclear and pixelated.

Information. The mean scores for the last section, Information, ranged from $\bar{x}=2.0$ (e.g. ShimSpine Exercises) to $\bar{x}=4.5$ (Vivira: Back, Knee, and Hip). In almost all apps, the accuracy of app description in the app stores was accurate. Moreover, the given visual information was clear, logical and correct in many apps. However, in this section, apps had not high credibility, with a majority of apps affiliated to a small NGO, institution or commercial business.

Scientific evidence and effectiveness

The search resulted in 5 apps (11.62%) supported by scientific literature, and 3 apps (6.97%) where the authors claimed a scientific background. 16 apps (37.20%) had no scientific background. 19 apps (44.18%) could not be specified because the email sent to the developer was left unanswered. Apps with scientific backgrounds are summarized in Appendix B.

Kaia. The study of Huber and colleagues (2017) was aimed to report retrospective short-term results of the Kaia-app regarding lower back pain treatment. Data of 180 users were evaluated for 12 weeks regarding the duration of use and effect on in-app user-reported pain levels using the numerical rating scale (NRS). Pain levels decreased from baseline NRS 4.8 to

3.75 for all users at the end of the observation period (Huber, Priebe, Baumann, Plidschun, Schiessl, & Tölle, 2017).

Spinal Risk Assessment Tool. In 2017, Veeravagu and colleagues conducted a prospective analysis of the novel Risk Assessment Tool for spinal surgery (RAT). The RAT (as a novel instrument for risk assessment) targeted patients undergoing spine surgery. It was developed based on administrative claims database. Data from patients were collected to compare the RAT to other surgical risk calculators, e.g. the ACS NSQIP. Results show that both tools were able to identify patient complications more likely after spine surgery. However, low accuracy in current measures was reported (Veeravagu et al., 2017).

SORT. Protopapa and colleagues (2014), aimed to develop and validate a preoperative risk stratification tool to predict 30-day mortality after surgery in adults. Logistic regression was used to construct a model in the derivation cohort to create the SORT, which was tested in the validation cohort. 16788 patients were analysed, with the SORT demonstrating better discrimination than another surgical risk scale. In conclusion, the SORT allows rapid simple data entry of six preoperative variables and provides a percentage mortality risk for individuals undergoing surgery (Protopapa, Simpson, Smith, & Moonesinghe, 2014).

SLIC. Short for the subaxial cercival spine injury classification system, the SLIC was developed by Vaccaro and colleagues (2007) to convey information about 3 categories: injury pattern, treatment considerations, and prognosis. The injury severity score is obtained by summing the scores from each of those categories. The SLIC suggested high construct validity. Reliability can be compared to other systems; making the SLIC a comprehensive classification system for subaxial cerival trauma (Vaccaro et al., 2007).

Spine Score. As a scoring and classification system for spinal pathology, Spine Score allows rapid access to information and calculation of clinical scores. The current range of scoring and classification systems within Spine Score include Myelopathy, Tumour, Infection, and

Trauma. The single classifications used in the app are all referenced on the website of the developer (Spine Score, 2020).

Apps claiming scientific background. When the emails addressing scientific evidence were sent to the developers, the main contributor of 'The Spine APP' answered, claiming the existence of links to PubMed listed publications. Unfortunately, these were nor provided neither found. Similarly, the developer of 'Deuk Spine Institute' claimed scientific background, which was not provided or found in databases or on their website. Another app, 'SpineScan3DR', claimed the validation of the compatible device as well as the validation of data collected and presented in the app. Nevertheless, results were missing and not yet published.

Mentionable Apps

Among all assessed applications, 'Vivira: Back, Knee, and Hip' reached the highest app quality mean score (4.76), followed by 'Kaia' (4.71) and 'Mineo' (4.67). The scores result from a high level of design on all 4 subscales of the MARS. Regarding the engagement and personalization, all 3 apps provided full customization regarding personal data, diagnosis, therapy goals, and exercise plans. Every component of the apps was aesthetically pleasant and well-explained, engaging the user to interact with the different sections of the app. The information displayed in the apps was highly relevant and featured tips and knowledge regarding the users' exercise plan, condition, and diagnosis, along with tips on how to master everyday life tasks. Examples of the personalization process, home screen and content of these apps shown in Appendix C.

In contrast, the apps with the lowest scores were 'Sciatic Nerve Pain Exercises' with an app quality mean of 2.05, and Shim Spine Exercises, with an app quality mean of 2.33. Both apps targeted exercise to help spinal surgery patients to recover. However, personalization was non-existent in both apps, with no opportunity of personalization or adjusting settings. The design of the app was somewhat confusing, with inconsistent elements, missing explanations on how to conduct exercises, with low-quality graphics and a lack of correct, well-written and comprehensive information. Examples of both apps can be found in Appendix E.

Assessed apps also include product-specific apps (Bauerfeind-App, a recovery-oriented exercise-based app which features a ligature or bandage), device-specific apps (SpineScan3D, features a device attached to the phone, measuring spinal deformities), assessment tools (e.g. assessing complications after surgery), and clinic-specific apps (‘INSELhealth’, supporting treatment/ recovery in the hospital). Design components as well as the content of those apps are summarized in Appendix F.

Discussion

Along with the rapid development and expansion of smartphone technologies, the increased use of mHealth in the medical sector holds great potential to optimize healthcare. This systematic review was conducted to address the question which examples of mHealth interventions are available for spinal surgery patients in the two leading app stores. Furthermore, the Mobile App Rating Scale was used to assess the content and quality of mHealth apps in conjunction with a corresponding scientific background. The search yielded 43 unique apps in total, mainly targeting physical health, the increase of well-being, exercise, and education. The overall app quality mean score was 3.64 out of 5, with 81.39% of the apps reaching the minimum acceptability score of 3.0. Regarding individual sections of the MARS, the highest mean score was recorded in Section B (Functionality) and the lowest in Section A (Engagement). The scientific evidence of the mHealth applications was sparse, with only 5 apps validated with scientific literature and 3 apps claiming a scientific background. This data suggests that the overall content and quality of most apps is acceptable, but they lack scientific evidence.

Interpretation

The results provide a broad overview of the general availability of mHealth apps targeted at spinal surgery patients, but also allow insight into the content, components, quality, and credibility of those apps. Regarding the scientific evidence of the apps, the results are in line with recent systematic reviews of similar health sectors evaluating mHealth apps using the Mobile App Rating Scale (Bahadori, Wainwright, & Ahmed, 2020; Creber et al., 2016; Machado et al., 2016). For example, the systematic app reviews of Creber et al. (2016), Bahadori et al. (2020) and Machado et al. (2016) reported similar app quality mean scores, as well as the highest mean scores in Functionality and lowest in Engagement. In addition to these similarities, all authors criticize the lack of evidence in mHealth applications, low app quality and missing medical accuracy. For example, the systematic review of Bahadori and colleagues (2020) examined smartphone apps for hip and knee replacement surgery patients. They reported that none of the examined apps was scientifically tested, underlining the need for medical accurate mHealth apps without commercial bias (Bahadori et al., 2020). Another recent systematic review of secure apps for daily clinical use conducted by German orthopaedic surgeons pointed out that the evidence and trustworthiness of those apps is strongly limited (Dittrich et al., 2020). This also corresponds to earlier conducted systematic reviews reporting the lack of evidence in mHealth apps, missing medical verification, validity, and app quality (Creber et al., 2016; Adam, Hellig, Perera, Bolton, & Lawrentschuk, 2018; Machado et al., 2016). An explanation for this phenomenon is that there are no existing rules regarding content, medical accuracy, and effectiveness of the apps, with no consensus on how to design and evaluate those apps (Dawson et al., 2019). Although the MARS is already a widely used, expert-based, and reliable method for assessing the quality of mHealth apps, it is used to assess already published apps. To ensure a good app design, as well as a good scientific basis of apps, guidelines should be developed to ensure quality enhancement in mHealth apps.

Implications

These systematic reviews, as well as the results of the current review, underline the issue of a weak evidence base, a lack of app testing and evaluation, and adequate app development guidelines. The findings highlight that there is still little known about the effectiveness of those apps and that apps designed for the medical sector need to be tested regarding their effectiveness. Consequently, without stated medical involvement, published apps could be mistaken by end-users as ‘professional and medical’. Patients could use those apps complementary to their traditional treatment, which may put them at risk. The danger is that without a medical background, validation or scientific testing, end-users with severe medical conditions could be harmed due to unknown outcomes of the app.

To prevent this issue, this research furthermore can serve as a guide for future developers, providing a basic overview of the design component, app content and quality. While previous studies using the Mobile App Rating Scale mainly focused on other sections in healthcare, this is the first systematic review focusing solely on spinal surgery patients providing a broad overview over a widely unknown field of mobile healthcare. It provides implications for content and design improvements, for example the need for more engagement in end users in terms of fun, degree of interest, customization, interactivity, along with the usage of more high-quality information along with the incorporation of medical professionals and credible sources. Moreover, it introduces already successfully implemented and well-made apps which could serve as an example for designing new mHealth technology related to spinal surgery patients.

Limitations

Although the findings of this study might be promising regarding a successful implementation in clinical practice, they should be handled with caution since the generalizability of the results is limited to apps targeting spinal surgery patients only. In addition, it remains questionable whether the app is generalizable to older people who may be more prone to spinal surgery interventions. Although most app descriptions claim the suitability for all age groups, older people often lack knowledge about how to use technological devices, which was not considered during the MARS evaluation. In addition, Briede-Westermeyer and his colleagues (2020) found a lack of interest and knowledge about how to use technological devices and their functions; along with the fact that only 39.8% of their participants aged >60 possessed a mobile phone. This is a noticeable increase compared to the 20% found by the study of Navabi and colleagues (2016). However, the number of elderly people using mobile phones might be rising, it remains questionable if the current user interface and design is suitable for older people since a digital native provided this review. Aside already existing health complications of older people, difficulties arise from normal progressive decline of physical and cognitive abilities, which could impair the mobile phone usage for them (Briede-Westermeyer et al., 2020). In conclusion, attempts should be made to a targeted mHealth UI- design for elderly people.

Another limitation is the validity of the data, which is impacted by the conduction of the review by 1 author only. This could cause bias within the qualitative content evaluation as well as within the MARS assessment, making the results and data somewhat subjective. Another issue is that some apps were only assessed using demo versions or screenshots. Apps requiring access were sometimes not fully accessible because of privacy reasons or hospital policies. Demo versions and additional screenshots were assessed instead, leading to potential undetected features and bias of data due to limited user experience. Another point is that apps were displayed differently depending on the device used. For example, some apps were not fully optimized for iPad, even though they were compatible with it. Moreover, apps that were designed for wearables could not be fully assessed due to missing wearables.

Recommendations

Regarding the limited evidence and the effectiveness of mHealth applications, better, long-term evidence is needed. Although the MARS is recognized as an excellent method for mHealth quality assessment, it addresses existing scientific evidence only in 1 item, which is usually summarized into sections' mean score (Dittrich et al., 2020; Stoyanov et al., 2015). Therefore, to make mHealth technologies more effective, guidelines are needed to design and evaluate mHealth apps. This includes not only transparency regarding developers and authors, the app itself, in-built functions, and features, but also underlying theoretical backgrounds and reliability. A solution could be to expand the MARS, adding a section addressing scientific evidence and effectiveness of the apps. In addition, a guideline for publishing medical apps should be developed to address these issues before publishing mHealth apps. Therefore, further studies should approach the development of design guidelines to evaluate medical mHealth applications and test those apps regarding their effectiveness.

Conclusion

This systematic review aimed to identify mHealth apps targeted at spinal surgery patients identifying their content, quality, and scientific background. Apps were assessed using the Mobile App Rating Scale (MARS). While the content of apps is mainly focused on physical health, the overall in-app engagement of users remained poor. Furthermore, the results indicate inadequate literature support of mHealth apps, leading to the conclusion that most mHealth apps are still lacking scientific corroboration. To ensure a successful implementation of mHealth apps in practice, usability and effectiveness testing is mandatory. Therefore, further research is needed to address scientific backgrounds and effectiveness of apps, to provide future patients with the best care possible and clear the barriers for a new way of delivering healthcare.

References

Adam, A., Hellig, J. C., Perera, M., Bolton, D., & Lawrentschuk, N. (2018). 'Prostate Cancer Risk Calculator'mobile applications (Apps): a systematic review and scoring using the validated user version of the Mobile Application Rating Scale (uMARS). *World journal of urology*, *36*(4), 565-573. doi: https://doi.org/10.1007/s00345-017-2150-1

Andersson, G., & Watkins-Castillo, S. I. (2014). Spine Procedures. Retrieved February 25, 2020, from https://www.boneandjointburden.org/2014-report/iie0/spine-procedures

Bahadori, S., Wainwright, T. W., & Ahmed, O. H. (2020). Smartphone apps for total hip replacement and total knee replacement surgery patients: a systematic review. *Disability and Rehabilitation*, *42*(7), 983-988. doi: https://doi.org/10.1080/09638288.2018.1514661

Bendixen, R. M., Fairman, A. D., Karavolis, M., Sullivan, C., & Parmanto, B. (2017). A user-centered approach: understanding client and caregiver needs and preferences in the development of mHealth apps for self-management. *JMIR mHealth and uHealth*, *5*(9), e141.

Briede-Westermeyer, J. C., Pacheco-Blanco, B., Luzardo-Briceño, M., & Pérez-Villalobos, C. (2020). Mobile Phone Use by the Elderly: Relationship between Usability, Social Activity, and the Environment. *Sustainability*, *12*(7), 2690. doi: https://doi.org/10.3390/su12072690

Creber, R. M. 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 and uHealth*, *4*(2), e74. doi: https://doi.org/10.2196/mhealth.5882

Dawson, R. M., Felder, T. M., Donevant, S. B., McDonnell, K. K., Card III, E. B., King, C. C., & Heiney, S. P. (2020). What makes a good health 'app'? Identifying the strengths and limitations of existing mobile application evaluation tools. *Nursing Inquiry*, *27*(2), e12333. doi: https://doi.org/10.1111/nin.12333

Debono, B., Bousquet, P., Sabatier, P., Plas, J. Y., Lescure, J. P., & Hamel, O. (2016). Postoperative monitoring with a mobile application after ambulatory lumbar discectomy: an effective tool for spine surgeons. *European Spine Journal*, *25*(11), 3536-3542. doi: https://doi.org/10.1007/s00586-016-4680-4

Dittrich, F., Beck, S., Harren, A. K., Reinecke, F., Serong, S., Jung, J., ... & Landgraeber, S. (2020). Analysis of Secure Apps for Daily Clinical Use by German Orthopedic Surgeons: Searching for the" Needle in a Haystack". *JMIR mHealth and uHealth*, *8*(5), e17085. doi: https://doi.org/10.2196/17085

Hou, J., Yang, R., Yang, Y., Tang, Y., Deng, H., Chen, Z., ... & Shen, H. (2019). The Effectiveness and Safety of Utilizing Mobile Phone–Based Programs for Rehabilitation After Lumbar Spinal Surgery: Multicenter, Prospective Randomized Controlled Trial. *JMIR mHealth and uHealth*, *7*(2), e10201. doi: https://doi.org/10.2196/10201

Huber, S., Priebe, J. A., Baumann, K. M., Plidschun, A., Schiessl, C., & Tölle, T. R. (2017). Treatment of low back pain with a digital multidisciplinary pain treatment app: short-term results. *JMIR rehabilitation and assistive technologies*, *4*(2), e11. doi:10.2196/rehab.9032

Istepanian, R. S., & Woodward, B. (2016). *M-health: Fundamentals and Applications*. John Wiley & Sons.

Jake-Schoffman, D. E., Silfee, V. J., Waring, M. E., Boudreaux, E. D., Sadasivam, R. S., Mullen, S. P., ... & Pagoto, S. L. (2017). Methods for evaluating the content, usability, and efficacy of commercial mobile health apps. *JMIR mHealth and uHealth*, *5*(12), e190.

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 Practice & Research Clinical Rheumatology*, *30*(6), 1098 1109. doi: https://doi.org/10.1016/j.berh.2017.04.002

Malvey, D., & Slovensky, D. J. (2014). *mHealth: transforming healthcare*. Springer.

Martin, T. (2012). Assessing mHealth: opportunities and barriers to patient engagement. *Journal of health care for the poor and underserved*, *23*(3), 935-941.

Mendiola, M. F., Kalnicki, M., & Lindenauer, S. (2015). Valuable features in mobile health apps for patients and consumers: content analysis of apps and user ratings. *JMIR mHealth and uHealth*, *3*(2), e40.

Moss, R. J., Süle, A., & Kohl, S. (2019). eHealth and mHealth. *European Journal of Hospital Pharmacy*, *26*(1), 57-58. Doi: http://dx.doi.org/10.1136/ejhpharm-2018-001819

Navabi, N., Ghaffari, F., & Jannat-Alipoor, Z. (2016). Older adults' attitudes and barriers toward the use of mobile phones. *Clinical interventions in aging*, *11*, 1371. doi: https://dx.doi.org/10.2147%2FCIA.S112893

Oh, H., Rizo, C., Enkin, M., & Jadad, A. (2005). What is eHealth (3): a systematic review of published definitions. *Journal of medical Internet research*, *7*(1), e1. doi: https://doi.org/10.2196/jmir.7.1.e1

Pohl, M. (2017). mHealth App Economics 2017/2018: Current Status and Future Trends in Mobile Health.

Protopapa, K. L., Simpson, J. C., Smith, N. C. E., & Moonesinghe, S. R. (2014). Development and validation of the surgical outcome risk tool (SORT). *The British journal of surgery*, *101*(13), 1774.

Rajak, M., & Shaw, K. (2019). Evaluation and selection of mobile health (mHealth) applications using AHP and fuzzy TOPSIS. *Technology in Society*, *59*, 101186. doi: https://doi.org/10.1016/j.techsoc.2019.101186

Semple, J. L., Sharpe, S., Murnaghan, M. L., Theodoropoulos, J., & Metcalfe, K. A. (2015). Using a mobile app for monitoring post-operative quality of recovery of patients at home: a feasibility study. *JMIR mHealth and uHealth*, *3*(1), e18. doi: https://doi.org/10.2196/mhealth.3929

Spine Score. (2020). Scoring and classification systems for spinal pathology. Retrieved July 02, 2020, from https://www.spinescore.org/

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 and uHealth*, *3*(1), e27.

Vaccaro, A. R., Hulbert, R. J., Patel, A. A., Fisher, C., Dvorak, M., Lehman Jr, R. A., ... & Fehlings, M. (2007). The subaxial cervical spine injury classification system: a novel approach to recognize the importance of morphology, neurology, and integrity of the disco-ligamentous complex. *Spine*, *32*(21), 2365-2374.

Vaishnav, A. S., & McAnany, S. J. (2019). Future endeavors in ambulatory spine surgery. *Journal of Spine Surgery*, *5*(Suppl 2), S139. doi: https://dx.doi.org/10.21037%2Fjss.2019.09.20

Veeravagu, A., Li, A., Swinney, C., Tian, L., Moraff, A., Azad, T. D., ... & Shuer, L. (2017). Predicting complication risk in spine surgery: a prospective analysis of a novel risk assessment tool. *Journal of Neurosurgery: Spine*, *27*(1), 81-91. doi: https://doi.org/10.3171/2016.12.SPINE16969

van Gemert-Pijnen, L., Kelders, S. M., Kip, H., & Sanderman, R. (Eds.). (2018). *eHealth research, theory and development: a multi-disciplinary approach*. Routledge.

Waller, A., Forshaw, K., Carey, M., Robinson, S., Kerridge, R., Proietto, A., & Sanson

Fisher, R. (2015). Optimizing patient preparation and surgical experience using eHealth technology. *JMIR medical informatics*, *3*(3), e29. doi: https://doi.org/10.2196/medinform.4286

World Health Organization. (2019, April 17). eHealth. Retrieved March 1, 2020, from https://www.who.int/ehealth/en/

Appendix A

Summary of the collected features and in-app functions of the assessed applications

App Name	SNS	Multipurpose	Rewards	Medication	Reminders	PainAssess	Datash	Progress	Checklist	Motivation
Myrecovery	N	Y	N	N	Y	Y	Y	Y	Y	Y
Mineo	Y	N	Y	N	Y	Y	Y	Y	Y	Y
Rally Recover	Y	Y	N	N	N	N	Y	Y	Y	Y
Spine& Tests	N	N	N	N	N	Y	N	N	Y	N
myHealthtrack	Y	Y	N	N	Y	N	Y	Y	Y	N
Suxeed	N	Y	N	N	N	Y	Y	Y	Y	N
PostMed	Y	Y	N	Y	Y	Y	Y	Y	Y	N
INSELhealth	N	N	N	N	Y	N	N	N	Y	N
Back Doctor (Free)	N	N	N	N	N	N	N	Y	N	N
Spine health	N	N	N	N	Y	N	N	Y	N	N
Back pain exercises	N	N	N	N	N	N	N	N	N	N
Rehab my patient	N	Y	N	N	N	N	Y	N	N	N
Deuk Spine Institute	Y	N	N	N	N	N	N	N	N	N

App Name	SNS	Multipurpose	Rewards	Medication	Reminders	PainAssess	Datash	Progress	Checklist	Motivation
Herniated Disc Exercises	N	N	N	N	N	N	N	N	N	N
herniated disc exercises	N	N	N	N	N	N	N	N	N	N
Sciatic Nerve Pain Exercises	N	N	N	N	Y	N	N	N	N	N
Lower back pain relief exercises	N	N	N	N	N	N	N	N	N	N
Vivira: Back, Knee, and Hip	N	N	N	N	Y	Y	Y	Y	Y	N
Home Physio	N	N	N	N	N	N	N	N	N	N
Lumbar herniated disc exercises	N	N	N	N	N	N	N	N	N	N
iLog	N	N	N	Y	N	N	Y	Y	N	N
Spine Score	N	N	N	N	N	N	N	N	Y	N

App Name	SNS	Multipurpose	Rewards	Medication	Reminders	PainAssess	Datash	Progress	Checklist	Motivation
Neurochirurgie	N	N	N	N	N	N	N	N	Y	N
SLIC	N	Y	N	N	N	N	N	N	N	N
Kaia	N	N	Y	N	Y	Y	N	Y	Y	Y
SpineScan3DR	N	N	N	N	N	Y	Y	Y	N	N
BACS	N	N	N	N	N	Y	Y	Y	N	N
Risk Assessment Tool for Spine Surgery	N	N	N	N	N	Y	N	N	Y	N
Allina Health	Y	Y	N	Y	Y	Y	Y	Y	Y	N
MyChart	Y	Y	N	N	Y	Y	Y	Y	Y	N
SORT	N	N	N	N	N	Y	N	N	Y	N
Ready Surgery App	N	Y	N	N	N	N	N	Y	N	N
Motion Spine Institute	Y	N	N	N	N	N	N	N	N	N
ShimSpine Exercises	N	N	N	N	N	N	N	N	N	N
Median- App	N	Y	N	N	N	Y	Y	Y	Y	N
Caspar Health Apps	N	Y	Y	N	N	Y	Y	Y	N	Y

App Name	SNS	Multipurpose	Rewards	Medication	Reminders	PainAssess	Datash	Progress	Checklist	Motivation
Median- App	N	Y	N	N	N	Y	Y	Y	Y	N
Caspar Health Apps	N	Y	Y	N	N	Y	Y	Y	N	Y
Kur-Reha BR	N	Y	N	N	N	Y	Y	Y	Y	N
Bauerfeind Therapie App	Y	Y	N	N	Y	N	Y	N	N	N
CatchMyPain	Y	Y	N	Y	Y	Y	Y	Y	Y	N
Pocket Spine Doc	N	N	N	N	Y	Y	N	Y	Y	N
Post Operative Solutions	N	Y	N	N	Y	N	N	N	N	N

Appendix B. Scientific Evidence of mHealth Apps

Table 5

Apps with scientific support

App Name	Provided scientific evidence
Kaia	Huber, S., Priebe, J. A., Baumann, K. M., Plidschun, A., Schiessl, C., & Tölle, T. R. (2017). Treatment of low back pain with a digital multidisciplinary pain treatment app: short-term results. *JMIR rehabilitation and assistive technologies*, *4*(2), e11. Doi: http://dx.doi.org/10.2196/rehab.9032
Risk Assessment Tool for Spine Surgery Procedures	Veeravagu, A., Li, A., Swinney, C., Tian, L., Moraff, A., Azad, T. D., ... & Shuer, L. (2017). Predicting complication risk in spine surgery: a prospective analysis of a novel risk assessment tool. *Journal of Neurosurgery: Spine*, *27*(1), 81-91. Doi: https://doi.org/10.3171/2016.12.SPINE16969
Spine Score	The website of the app provides a section called ‘research papers’ which the app was based on: https://www.spinescore.org/research-papers
SLIC	Literature for this scale is referenced online and can be found here: http://blog.digitalneurosurgeon.com/ Vaccaro, A. R., Hulbert, R. J., Patel, A. A., Fisher, C., Dvorak, M., Lehman Jr, R. A., ... & Fehlings, M. (2007). The subaxial cervical spine injury classification system: a novel approach to recognize the

App Name	Provided scientific evidence
	importance of morphology, neurology, and integrity of the disco-ligamentous complex. *Spine*, *32*(21), 2365-2374.
SORT	Protopapa, K. L., Simpson, J. C., Smith, N. C. E., & Moonesinghe, S. R. (2014). Development and validation of the surgical outcome risk tool (SORT). *British Journal of Surgery*, *101*(13), 1774-1783. Doi: https://doi.org/10.1002/bjs.9638

Apps claiming scientific backgrounds

SpineScan3DR

In collaboration with the University of Hong Kong, the SpineScan3D was developed as a tool to detect back surface topographic changes in spinal deformities (SpineScan3D, 2020). The device collects tilt information in axial, coronal and sagittal planes. The information can be accessed via the app. Although the user report is still to be launched on the website, a protocol for validation testing10 subjects is provided. However, results are nowhere to be found. All information can be found at the developers' website: http://www.spinescan3d.com/

The Spine App and Deuk Spine Institute.

As already mentioned, developers of the app were contacted and asked whether the app has been tested or verified by evidence in published scientific literature. The main contributor of the 'The Spine App', Marin Guentchev, answered

‘*Thanks for your email. The effect of the app has not been studied so far. The goal of the app is to provide evidence-based knowledge to spine patients. Thus we have links to PubMed listed publications.*’

Consequently, the developers’ website (https://thespineapp.com/en/index.php) was accessed and searched for links, but those were nowhere to find. Similarly, this happened with the developer of the ‘Deuk Spine Institute’ app, claiming that the app was developed based on scientific literature. But on their website, (https://deukspine.com/about-us/deuk-spine-foundation/) nothing could be found.

Appendix C. Examples of well-designed applications

1.0 Examples of the App Personalization Process

1.1 Mineo

Figure 2. First Part of the Personalization Process: ‘Mineo’

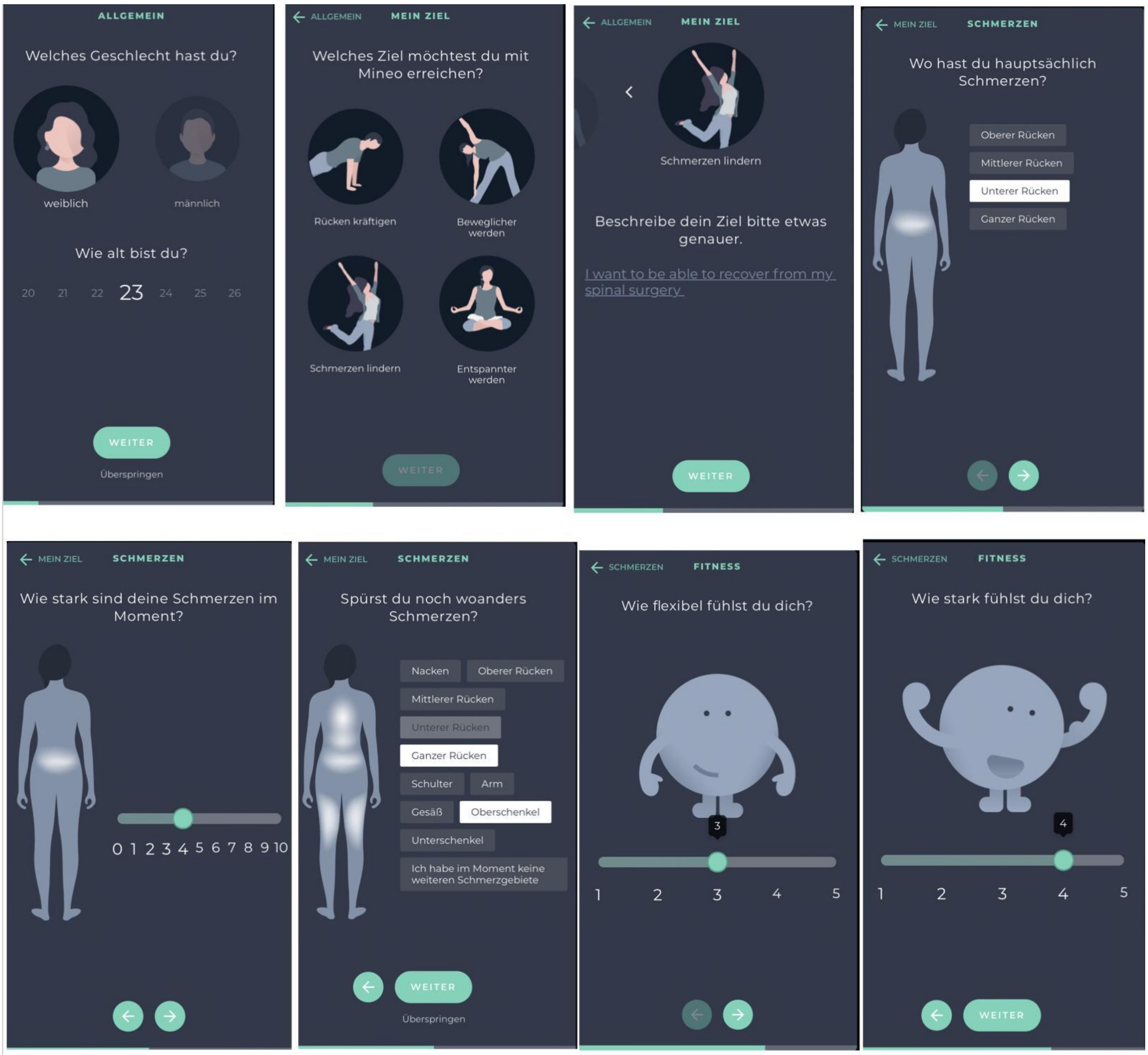

Figure 2. The screenshots show the personalization process the user encounters when opening the app for the first time. First, the user has to enter his gender and his age on a scale. Second, a goal assessment is made, asking the user what goal he wants to reach out of 4 provided pre-sets: ‘strengthening the back’, ‘get more flexible’, ‘reduce pain’, ‘get more relaxed’. After that, the user can enter a more specific goal for himself, here displayed as ‘*I want to be able to recover from my spinal surgery’*. In addition, the user can determine the area of his pain, followed by two 5-point scales, asking for the user’s flexibility and strength.

Figure 3. Second part of the Personalization Process: 'Mineo'

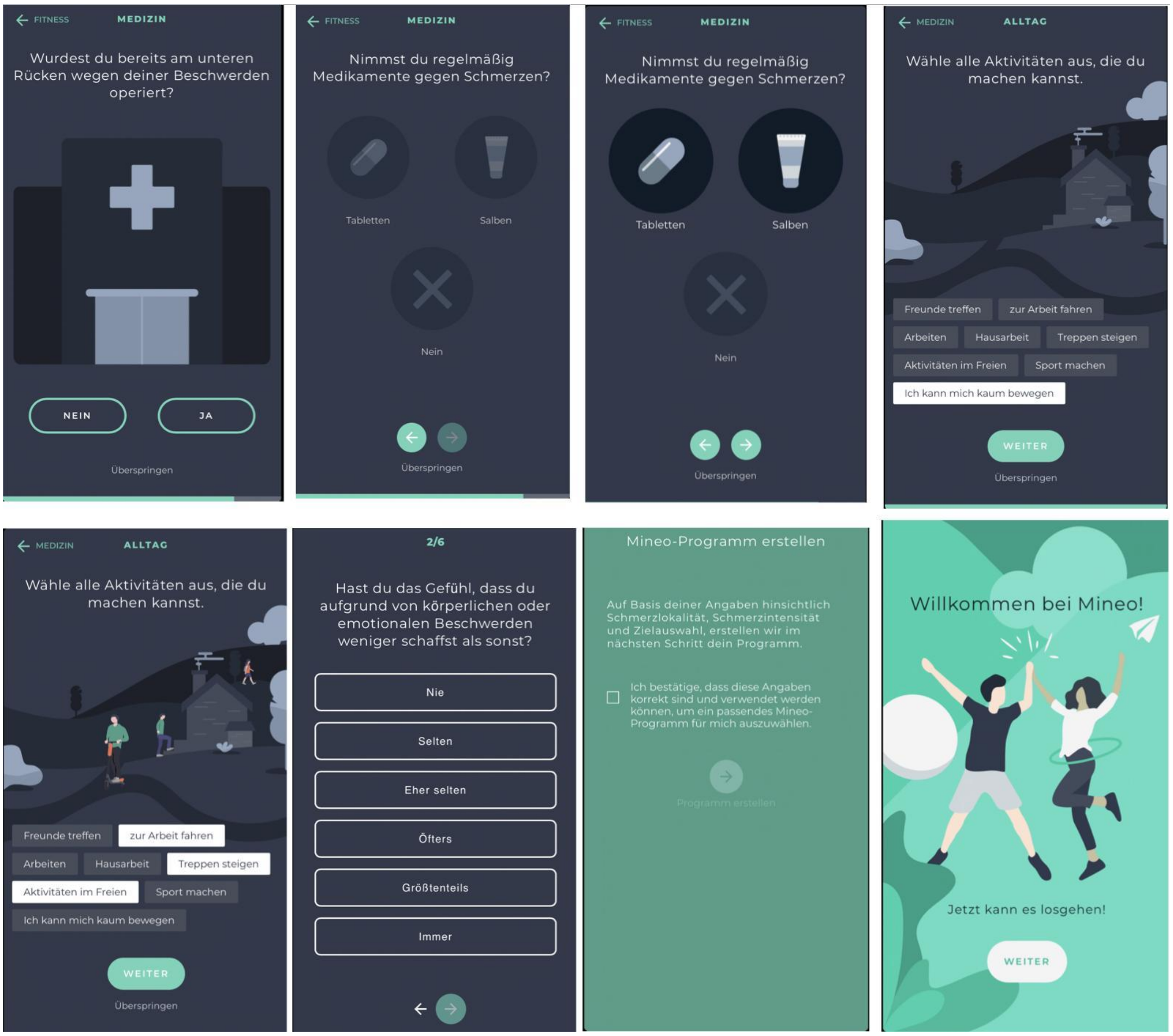

Figure 3. The screenshots above provide insight into the personalization process of the app. After the user answered some demographic questions regarding his gender, age, goals, areas of pain, etc., the user is asked if he already had surgery, and whether he takes any kind of medication. If yes, the names can be entered (and later, the user gets reminded to take them). Lastly, the user is asked which activities he can perform, ranging from 'performing exercise', or 'meeting friends', to 'not able to move at all'. At the end, a question is asked regarding emotional discomfort associated to the physical pain, before the app creates a personalized exercise plan.

1.2 Kaia

Figure 4. Personalization Process: Kaia

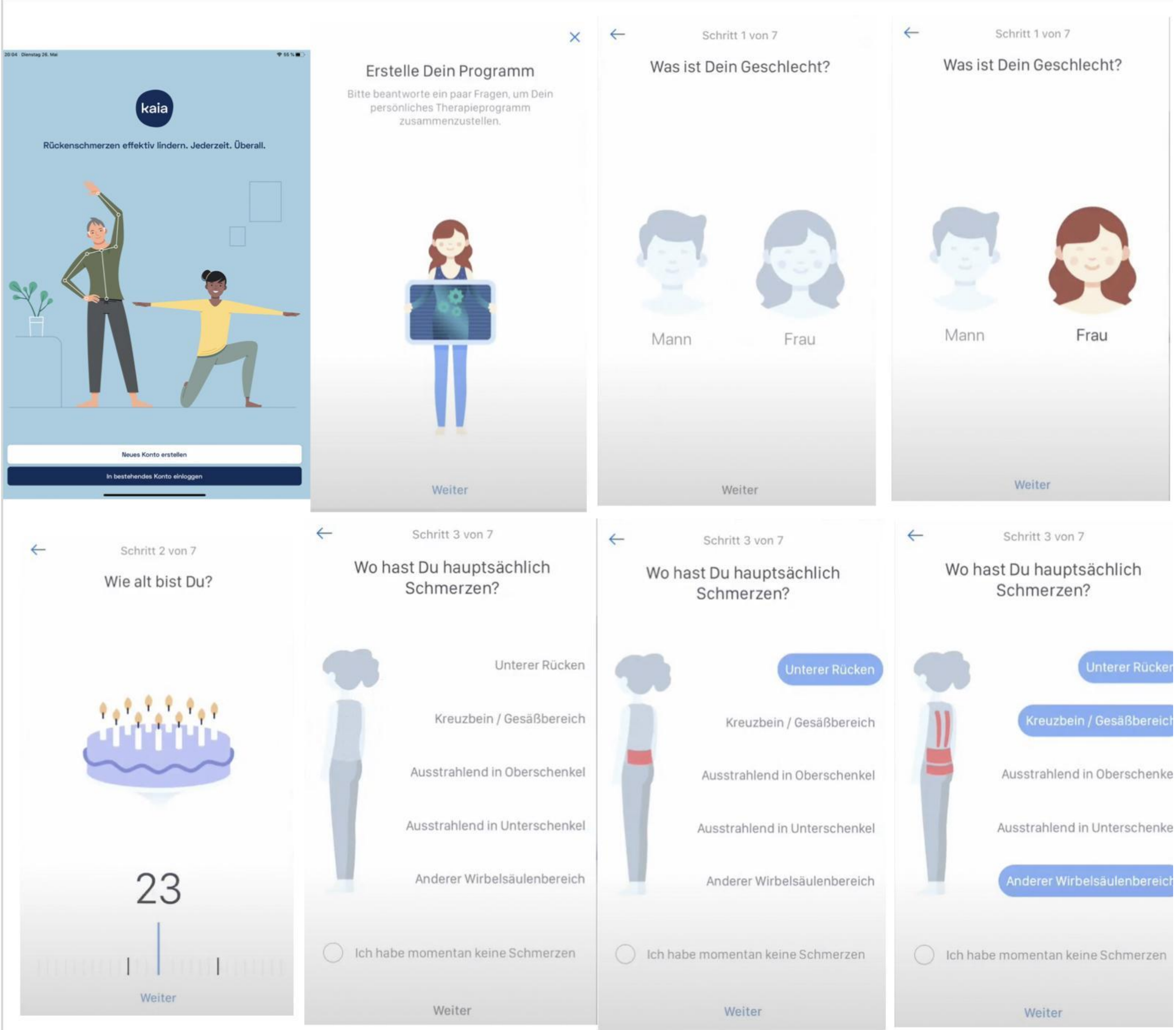

Figure 4. Personalization process the patient/user faces when opening the app for the first time. First, the user is required to register, using a code provided by healthcare providers or the user's health insurance. After that, a program is created based on the following information: age, gender and the area of pain. When the required information is entered, the app creates a personalized exercise program.

1.3 Vivira

Figure 5. Personalization process: Vivira

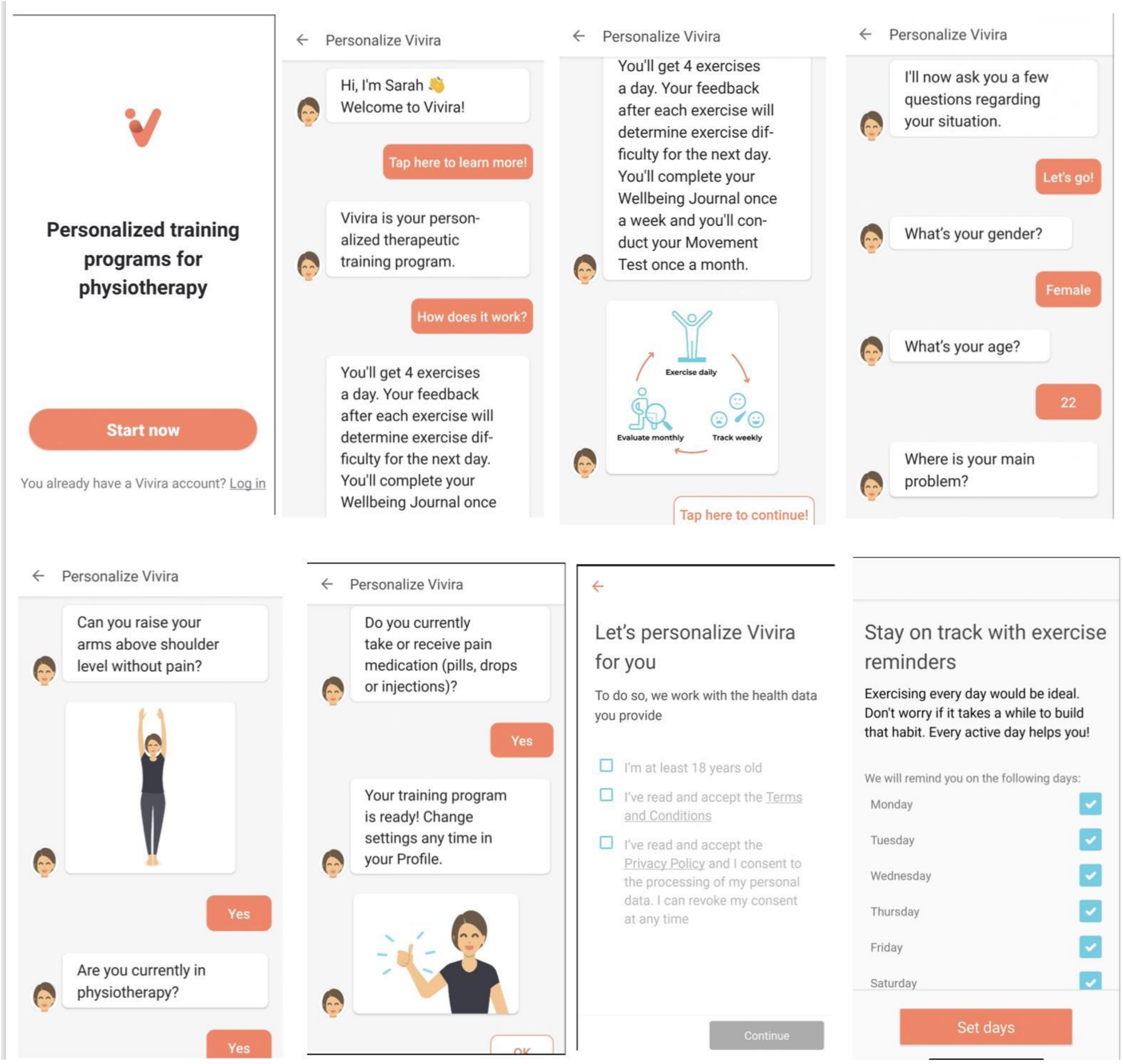

Figure 5. The personalization process of the app Vivira when the user opens the app for the first time. The user talks to a virtual assistant, providing the user with information on how the app works. Afterwards, demographic questions are asked, along with questions regarding the medical condition of the user. These are used to provide a broad diagnosis and to create a personalized exercise plan. Furthermore, the user gets the opportunity to schedule reminders.

2.0 Main Menu Examples

2.1 Mineo

Figure 6. Example of the Main Menu: Mineo

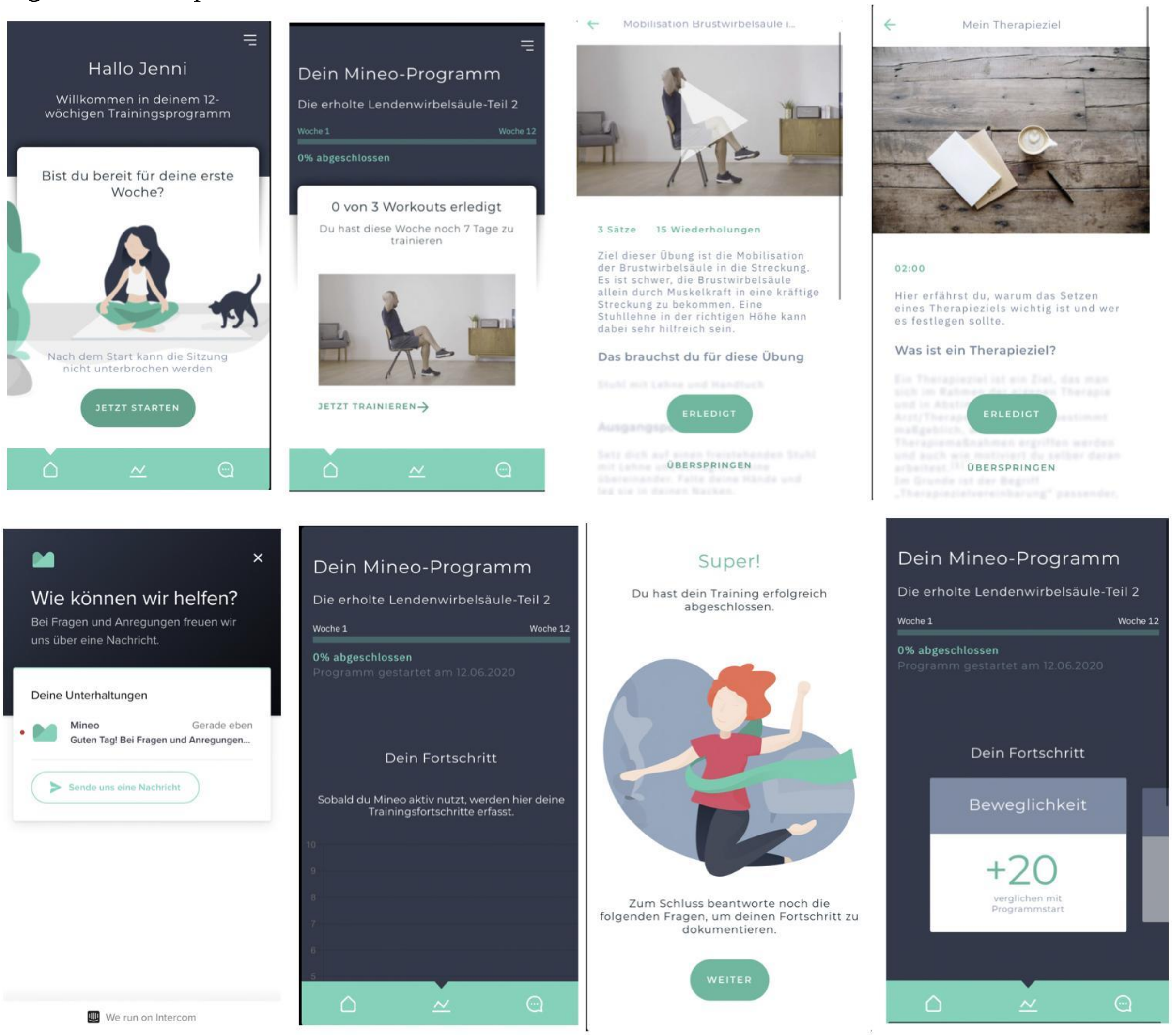

Figure 6. Example of the main menu. On the home screen (left bottom), the 12-week exercise program is shown. Exercises are well explained with a video and a person guiding through the exercises with a calming voice and a demonstration of the exercises. Before the user starts the video, a well-written explanation is shown, giving information about the purpose of the exercise, execution and needed materials. Afterwards (after the 1st training of the 1st week) the user can enter a personalized therapy goal. The app provides information why this is important for therapy success and commitment. When clicking on the little chat symbol in the right bottom corner, the user can send messages and ask therapy-related questions. When touching the symbol next to the chat symbol, the user can see and monitor his progress and gets points with every exercise he carries out. Furthermore, the user gets motivated when finishing the whole exercise for the day.

2.2 Kaia

Figure 7. Examples of the main menu: Kaia

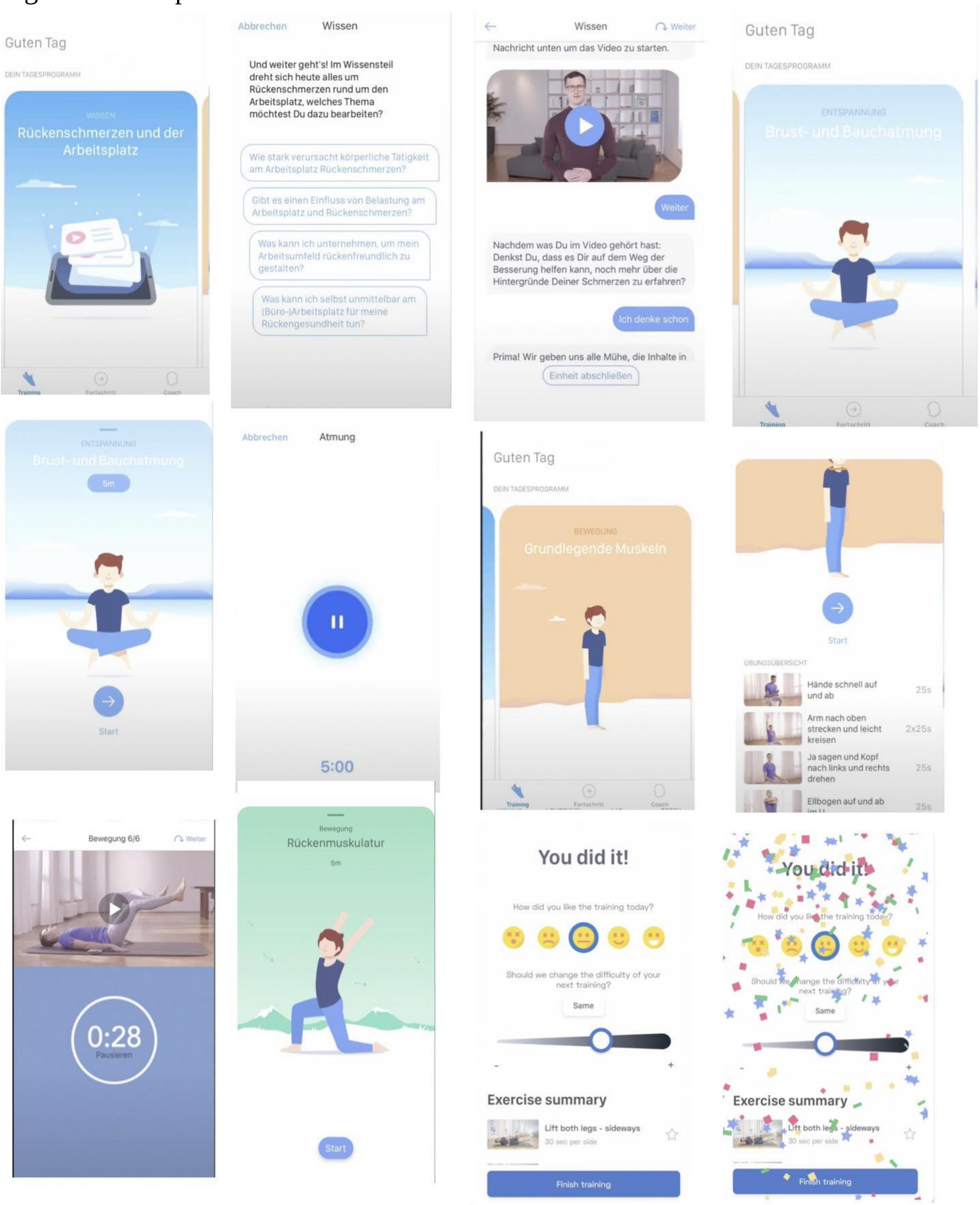

Figure 7. The main menu of Kaia features 3 different sections: ‘Wissen’ (Knowledge), ‘Entspannung’ (Relaxation), and ‘Bewegung’ (Movement), which content changes overtime. When entering the section ‘Wissen’, the user can read about daily changing topics, and give feedback whether the material is understood and clear. In the section ‘Entspannung’, the user can engage into guided meditations (voice). In the section ‘Bewegung’, the user can choose between daily changing workouts and receives motivation when finishing them. Furthermore, the user can give feedback.

2.3 Vivira

Figure 8. Examples of the main menu: Vivira

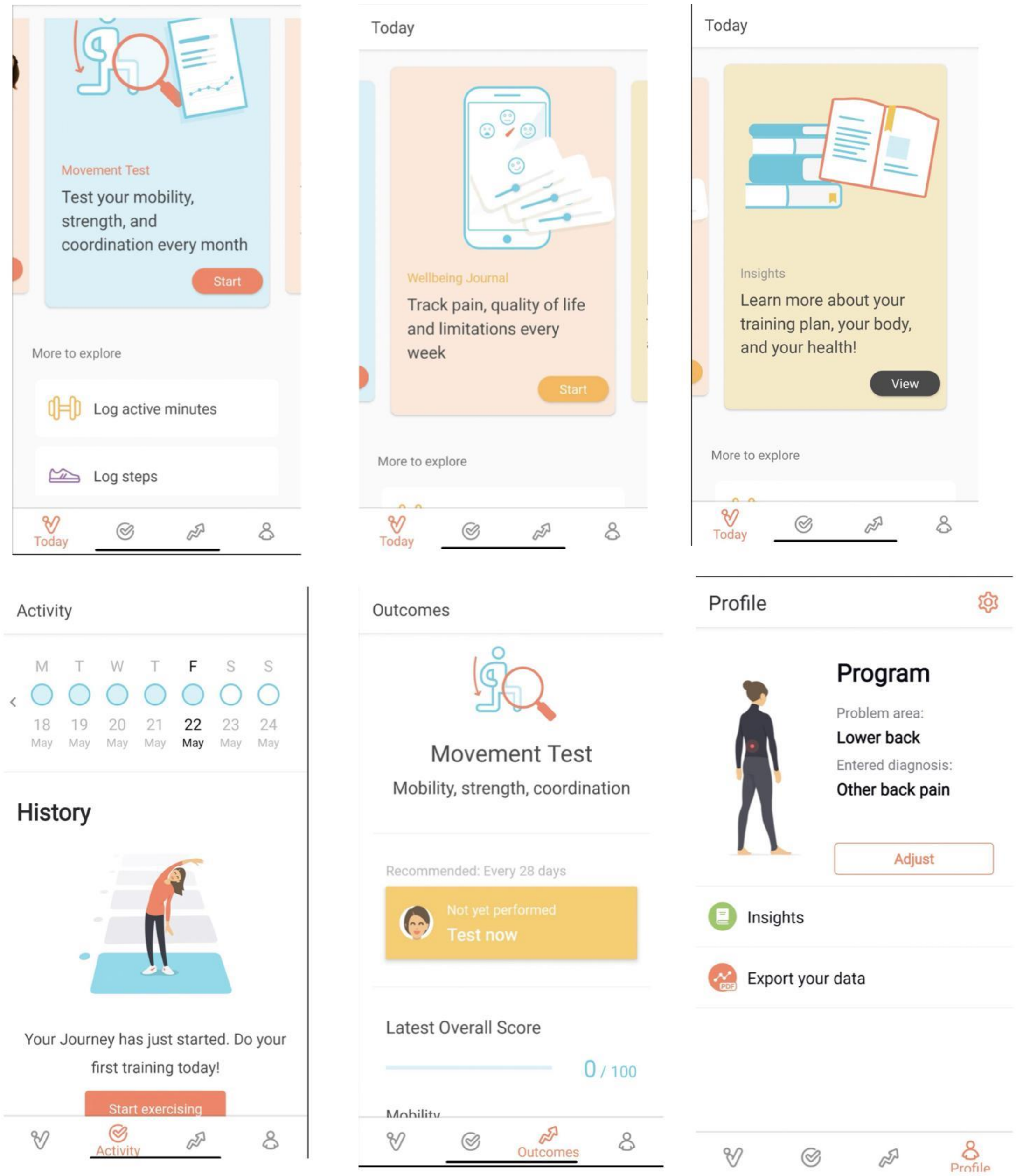

Figure 8. The main menu of Vivira contains 3 sections, providing the user with information regarding back pain and give the opportunity to assess pain, QoL, mobility, strength, and coordination. Furthermore, the user can assess health data like minutes exercises or steps taken throughout the day. In order to do that, the user can connect the app with a wearable. When accessing the ‘activity’ menu, the user can monitor his exercise history and progress. Under ‘profile’, insight is given into the personal data, problem area and entered diagnosis. Furthermore, the data can be exported and provided to an involved healthcare professional.

Appendix D- Examples of bad design

Figure 9. Sciatic nerve pain exercises

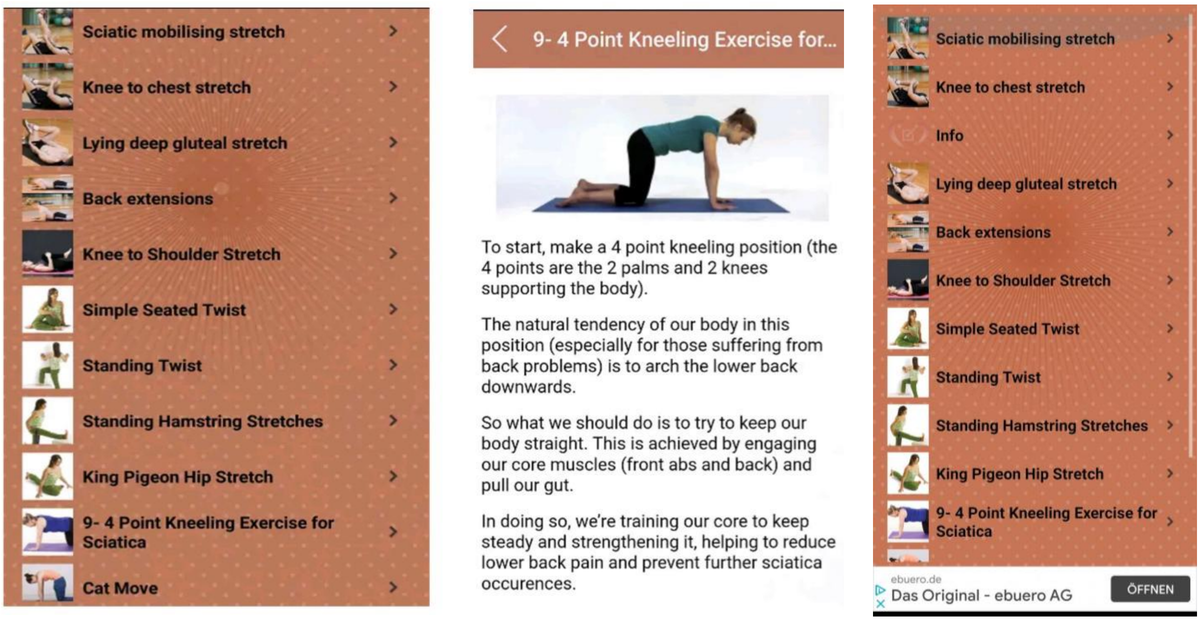

Figure 9. The home screen and exercise component of the app Sciatic Nerve Pain Exercises. When opening the app, the user directly sees the orange screen. The other (and only) available screen of the app are the displayed exercises, which are neither well-explained nor shown. There mostly is only one picture, which can be confusing and misleading, leading to a wrong execution of the exercises. In addition, sometimes advertisement is displayed as a banner at the top or bottom of the app.

Furthermore, the background is distracting and lowers the readability of the black letters. There is no banner with different menu points to click on, as well as a menu button. There are no means of personalization.

Figure 10. ShimSpine

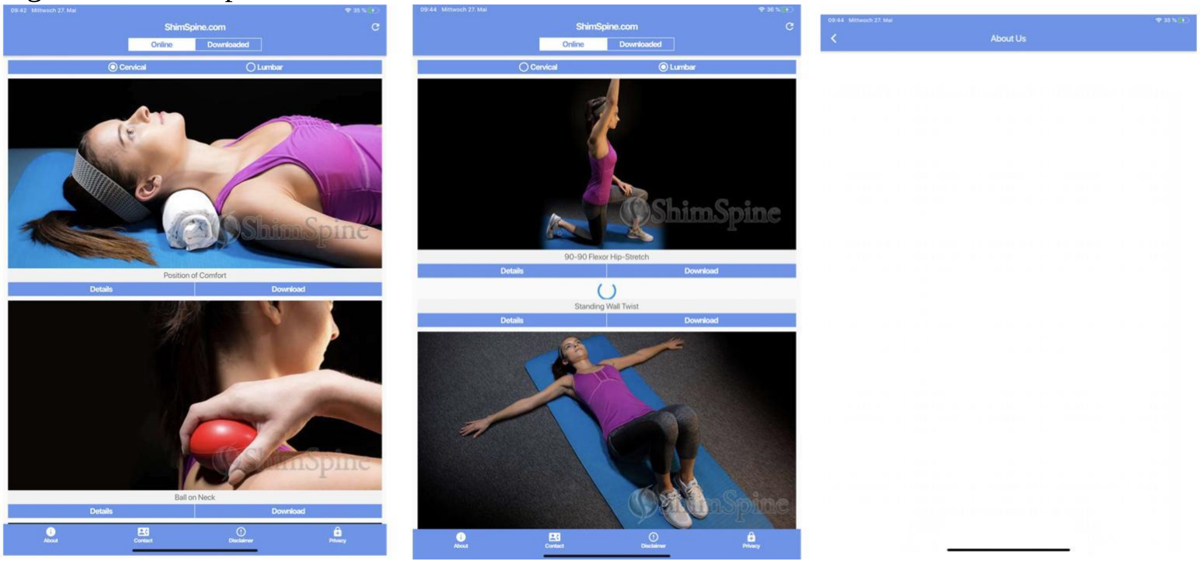

Figure 10. The main menu of the Shim Spine App. The user can choose between different exercise options: cervical and lumbar exercises. When clicking on those pictures, a small explanation shows up, describing the execution of the exercise insufficiently and in 1-2 sentences. The execution is mostly unclear and misleading. Furthermore, the ShimSpine logo is very interrupting when looking on the pictures. When clicking on the menu button on the left bottom (about us), only a white screen shows up with no information. Furthermore, it seems that some exercises are not displaced at all (and instead marked with a ‘loading’ circle), making access nearly impossible.

Appendix E- Categorical app examples

1.0 Apps bound to a specific product

Figure 11. Example of the Bauerfeind-App

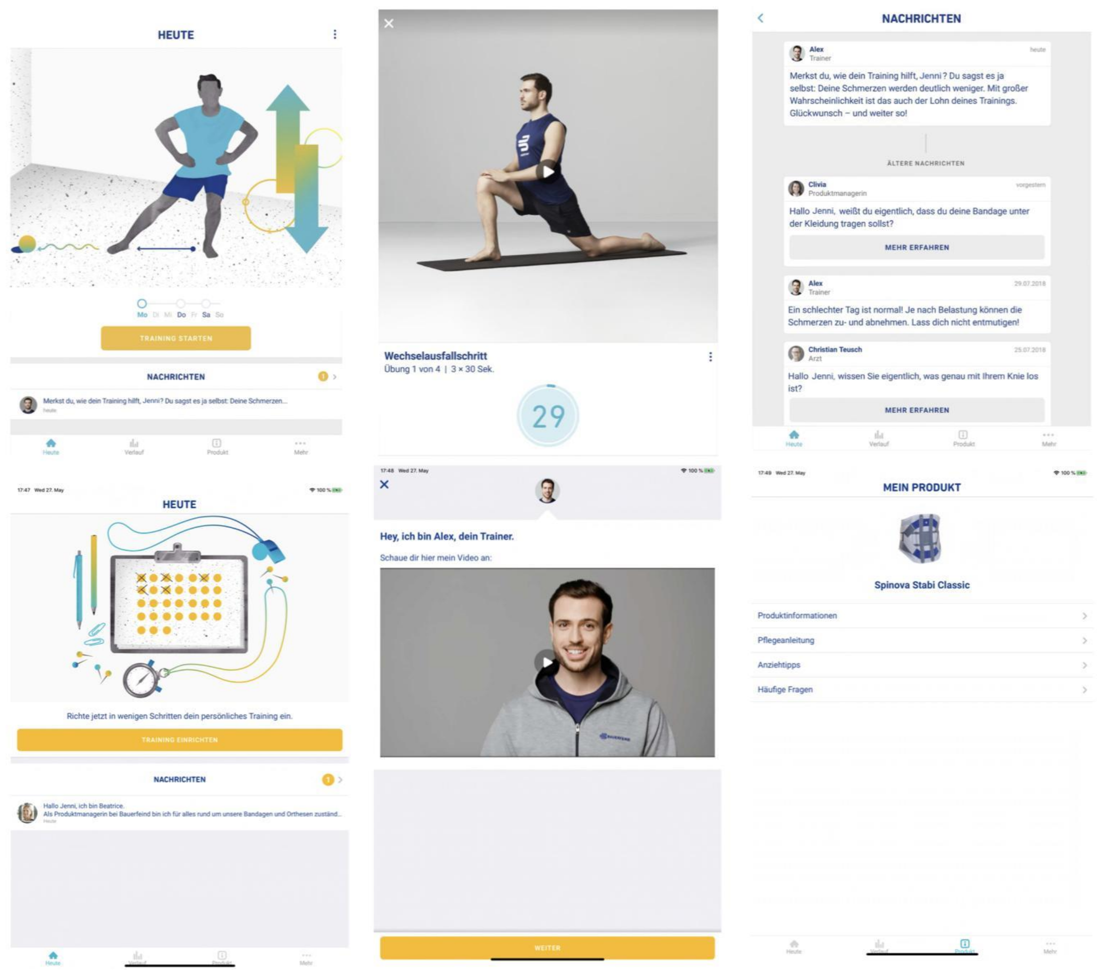

Figure 11. The main menu and functions of the Bauerfeind-app. When touching the home-button, the user can access his exercise plan and monitor the days when he must exercise. The exercise is well-explained and shown in a video by a personal trainer, which guides the user through the exercise. Furthermore, the user gets messages from the trainer and product manager and is able to ask questions. Moreover, a lot of information is provided regarding the product, pain, and exercise. When the user touches the ‘product’ button, information related to the product is given.

Figure 12. Function of the SpineScan3D Application

Figure 12. When entering the SpineScan 3D app, the user first needs to watch a video, which explains how the product works, how to attach the smartphone to it and how to record data.

First, the user needs to create an account to access the main menu. After that, the app automatically searches for the device, which is compatible via a Bluetooth connection. Once the device is synchronized, the user can start using the device. The collected data can be saved and is accessible via the main menu of the app.

2.0 Examples of Assessment tools

Figure 13. Examples of the SLIC-App

Figure 13. The main screens of the SLIC app. First, the user has to enter morphology, followed by disco-ligamentous complex information and neurological status. After touching the submit button, a new page 'results' shows up with the results and SLIC score. The app is based on scientific literature, which can be accessed through the app.

Figure 14. Example of the Risk Assessment Tool for Spine Surgery Application

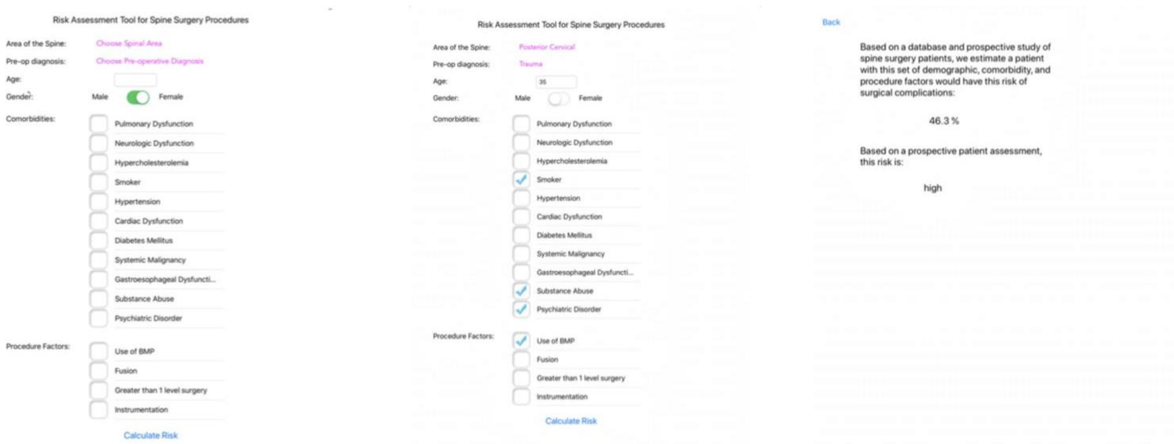

Figure 14. The main screens of the Risk Assessment Tool for Spinal Surgery Procedures application. When opening the app, the user has to enter information regarding the spinal area, pre-operation diagnosis, age, gender, conditions, and procedure factors. After touching the 'calculate risk' button, the risk of surgical complications is displayed in percent along with the accompanied risk. The app claims the use of scientific literature, which can be accessed through the app.

Figure 15. Example of the SORT Application

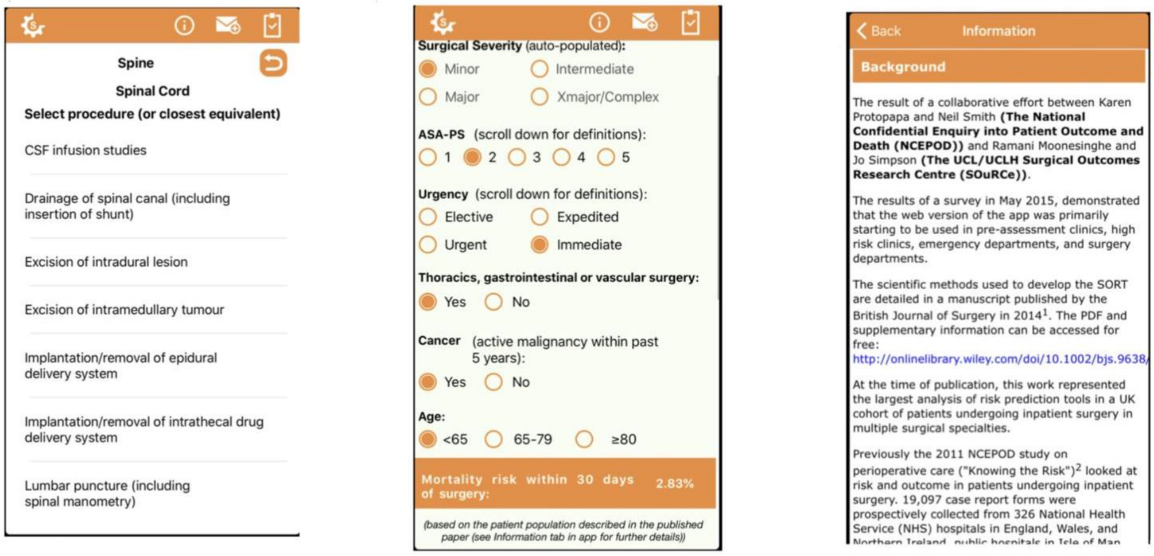

Figure 15. Example of the main screens of the SORT- Application. First, the user needs to answer several questions on different screens, before entering information regarding surgical severity, ASA-PS, Urgency and more. From this information, the mortality risk within 30 days of surgery is immediately calculated. When touching on the little ‘I’ button on the top of the screen, information can be accessed regarding the background and scientific method of the application.

3.0 Apps related to a specific clinic

Figure 16. Examples of the Clinic-Related Median-Application

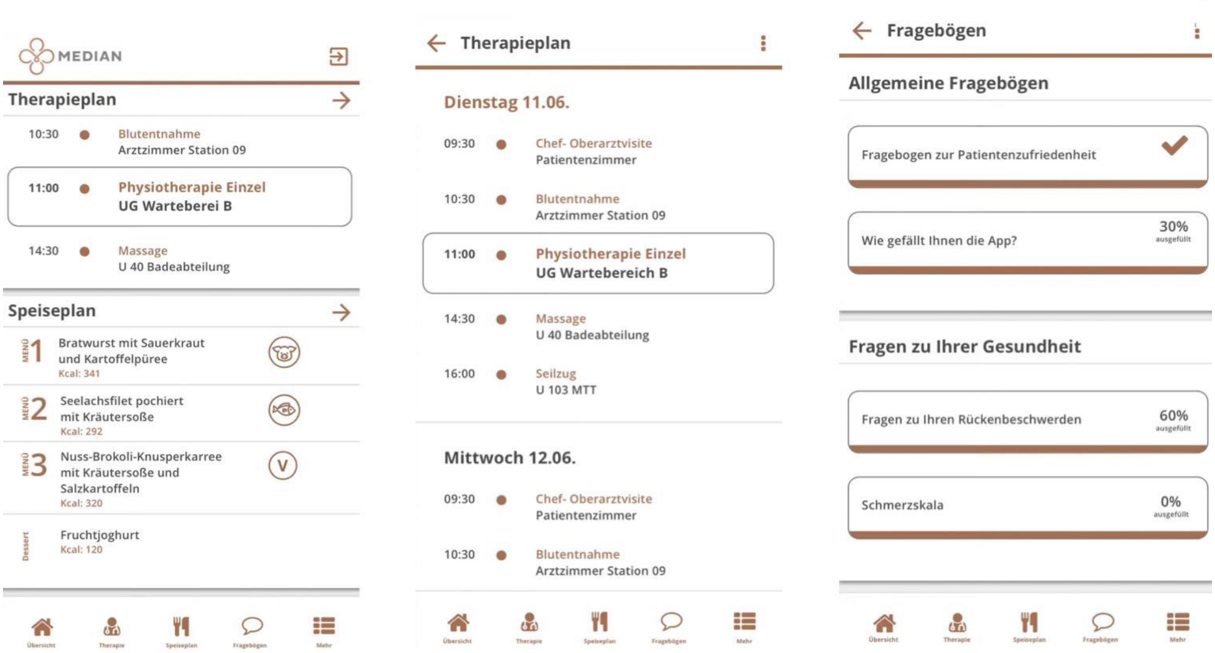

Figure 16. Main components of the Median- application. Through the single buttons on the menu, the local ‘Therapieplan’ (therapy schedule), ‘Speiseplan’ (meal plan), ‘Fragebögen’ (questionnaires) and other information is displayed, which is actualized in real-time. When accessing the therapy plan, the user can see the kind of therapy, the time and location. On the section ‘Speiseplan’, the different menus for the day can be seen, along with calories and contents (meat, fish, vegetarian). When touching on the ‘Fragebögen’- button, questions are asked related to the patient satisfaction, back pain, pain scale, and satisfaction regarding the app.

Figure 17. Example of the App ‘Kur- Reha BR’

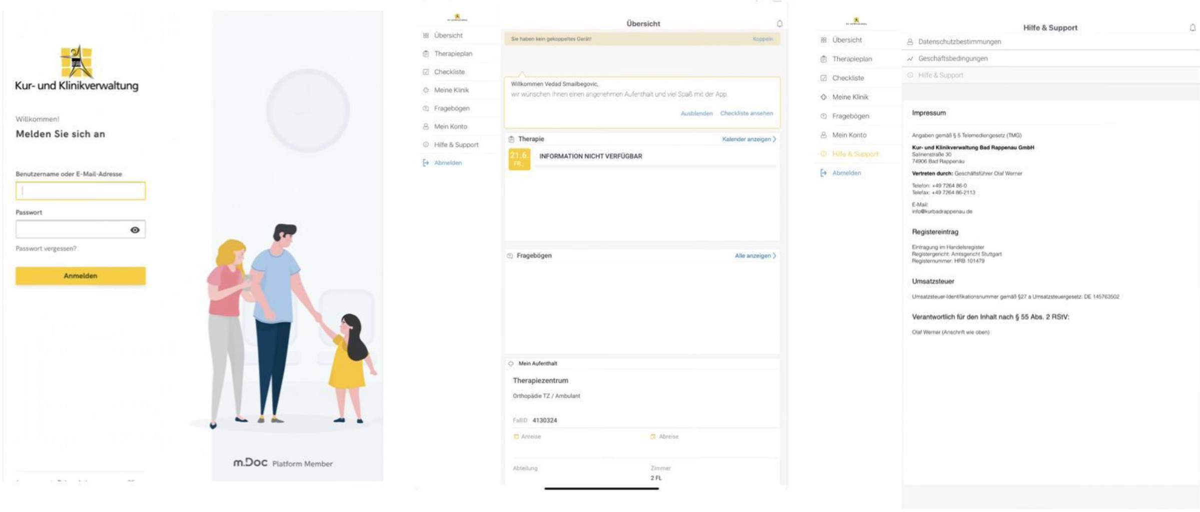

Figure 17. Main screens of the app Kur-Reha BR. When accessing the app for the first time, the user has to log in. After that, he can access the home screen. Categories are sorted in a bar on the left side of the app, where the user can access the therapy plan, checklists, information regarding the clinic, questionnaires, personal account information, and help& support.

Figure 18. Examples of Allina Health

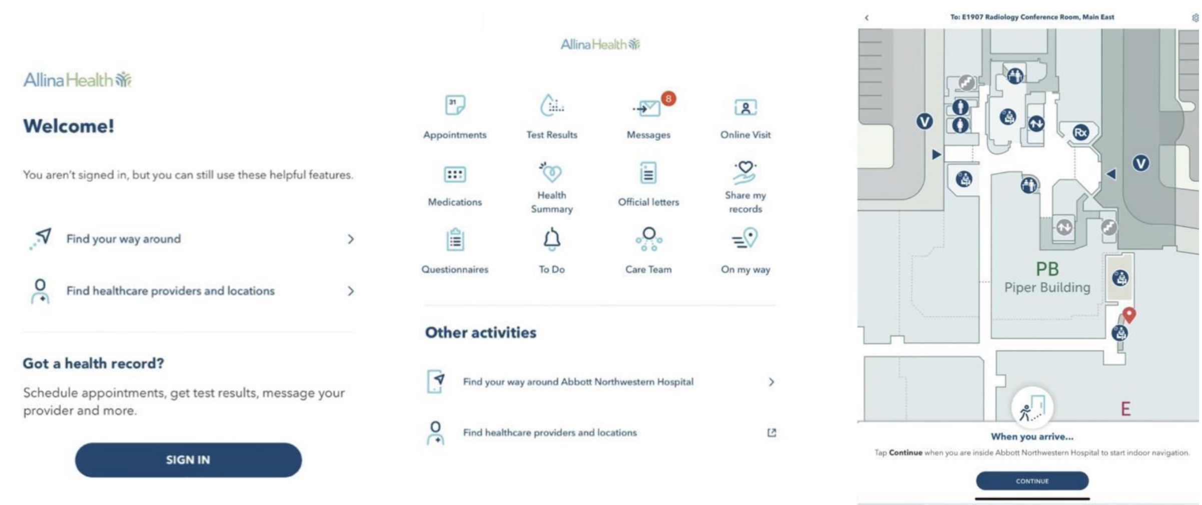

Figure 18. The main screen of Allina Health. When first accessing the app, the user has to sign in with a code provided by the clinic. When entering the main home screen, the user can manage appointments, see test results (which are synchronized into the app), see messages from healthcare providers and/or the clinic, can make an online visit, monitor medications, has insight into his health data, and many more. Below the main menu (other activities), you can also access a map from the hospital for orientation and finding the way to certain rooms.

Figure 19. INSELhealth

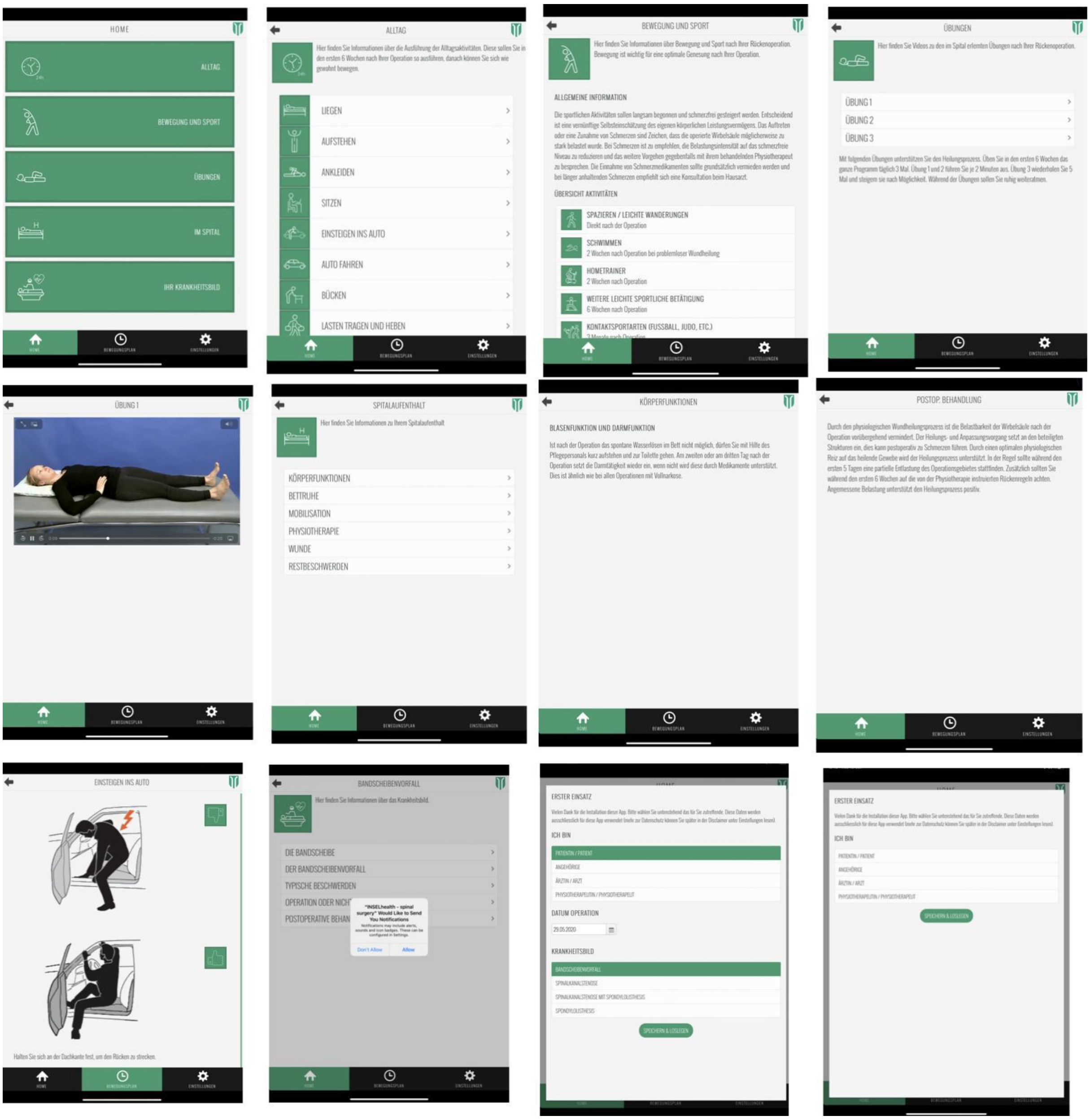

Figure 19. Screens and contents of the app INSELhealth. The two screenshots in the right bottom show the personalization process when first entering the app. The user has to enter whether he is a patient, a relative, a therapist or a healthcare professional (doctor). Furthermore, the date of the surgery needs to be entered, as well as the diagnosis (herniated (slipped) disc, spondolysis). After that, the user enters the home screen (top left corner), where different categories are shown: everyday life, movement and sport, exercises, hospital-related information, diagnosis. When touching the field ‘Alltag’ (everyday life), the user gets information on how to perform daily tasks correctly and safely. When accessing ‘Bewegung’ (movement and sport), the app shows which kind of sport is possible to execute after the surgery. When touching ‘Übungen’ (exercise), the user can access exercises personalized to him and his diagnosis. When touching the other

categories, lots of information is available regarding the patients' recovery phase. The app can also send daily reminders.

4.0 App with good pain and recovery monitoring

Figure 20. Examples of Progress Monitoring and Pain Assessment: 'my recovery'- app

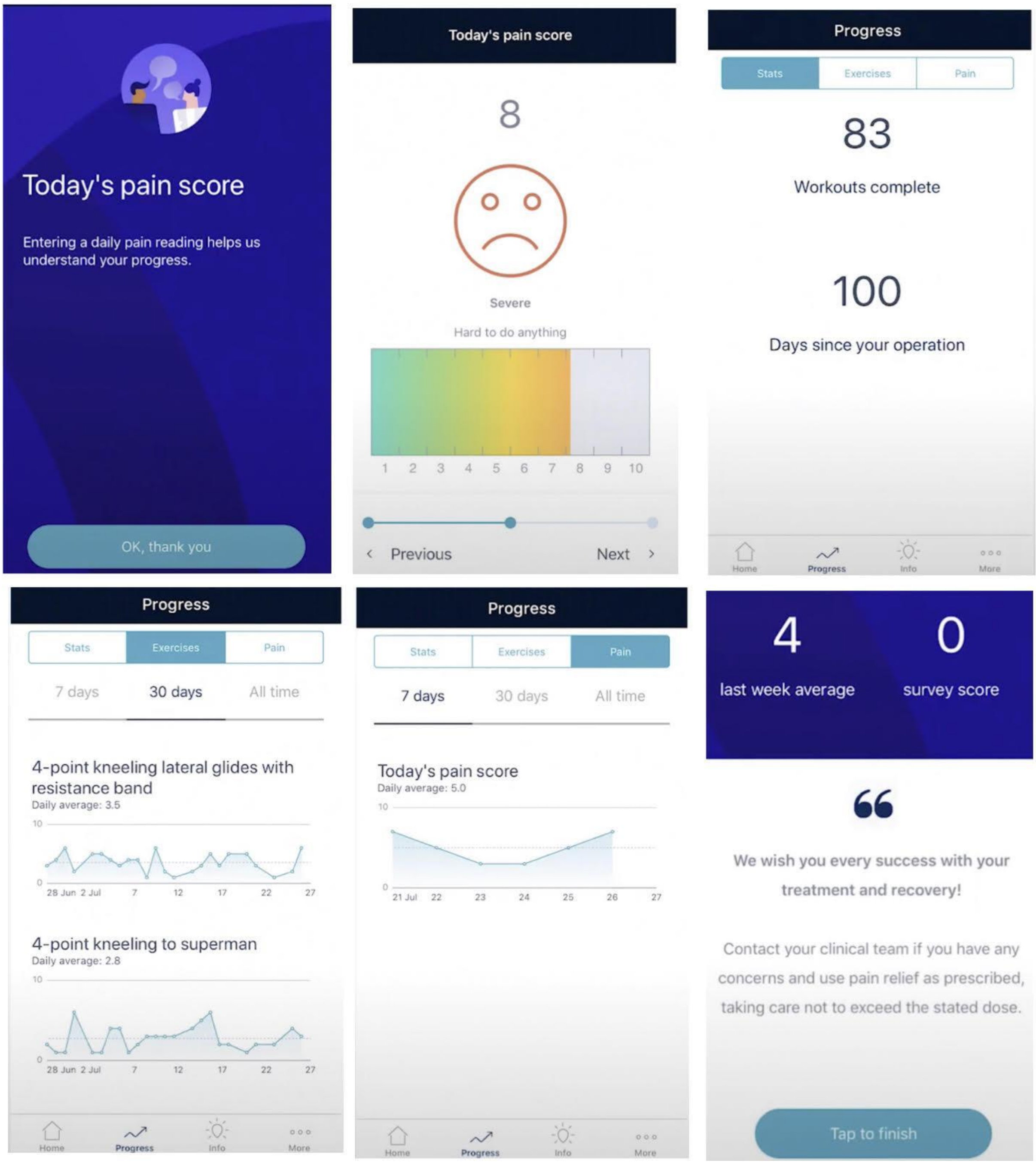

Figure 20. Examples of the progress-monitoring component and pain score assessment of the app 'my recovery'. The user has the opportunity to assess his pain with the aid of a pain scale.

Furthermore, the user can see his statistics regarding this pain score, as well as his exercise progress and statistics.

Figure 21. Examples of the app ‘CatchMyPain’

Figure 21. Main screens of the app ‘CatchMyPain’. The user is required to log in before accessing the main screens of the app. The user is able to enter his back pain, along with other information (Intensity, Description, Condition (Stress and Fatigue have to be purchased)). Through a graph, the user is able to monitor his pain and can also share his records within an in-built social network system.

Note: the screenshots are taken from the app store, since the pain monitoring and social network system can be shown in a better way. Furthermore, the medication system cannot be shown due to a server error during the periods the screenshots were taken.

5.0 Apps with Education and Information

Figure 22. Examples of the App ‘Pocket Spine Doc’

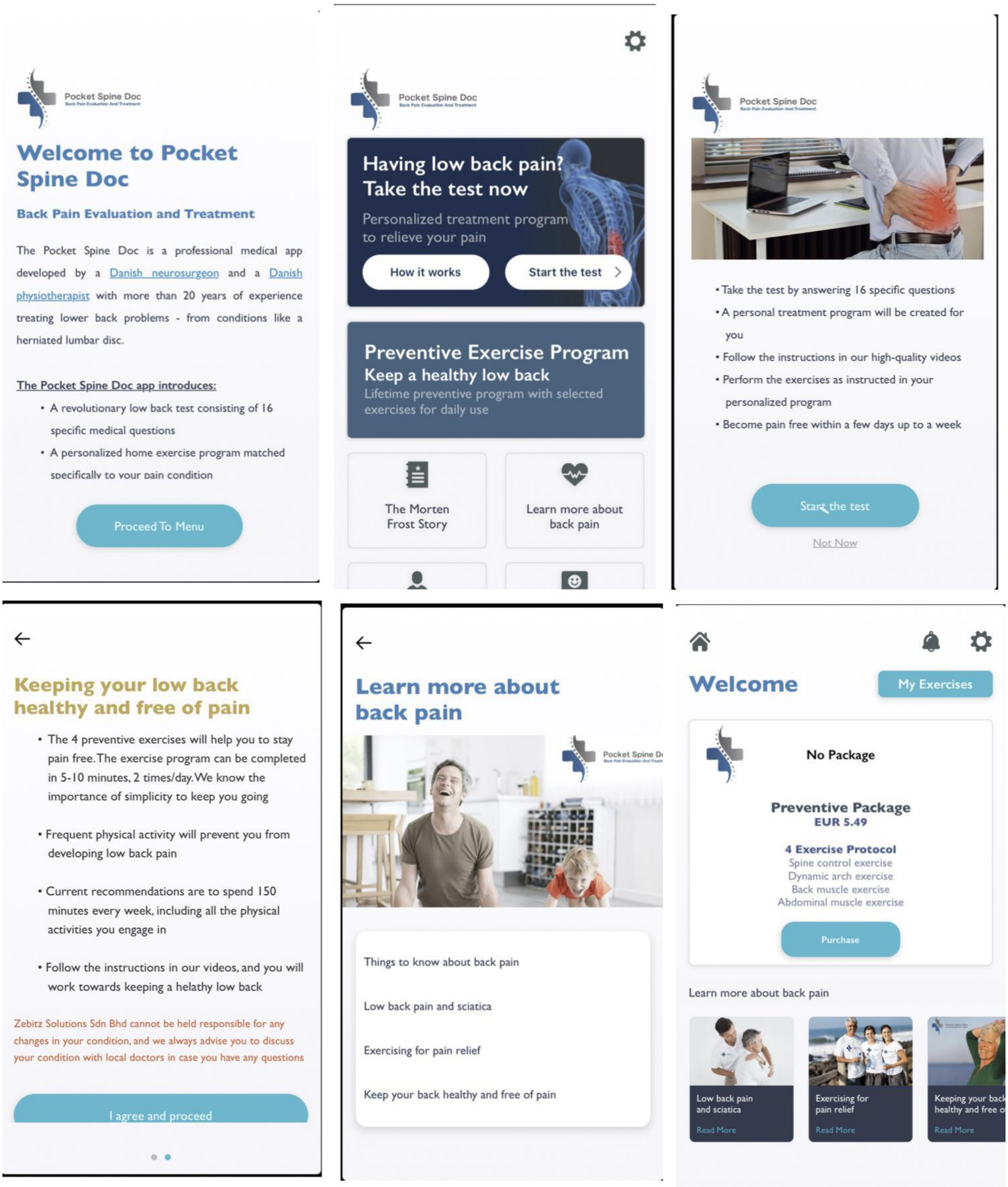

Figure 22. Example screens of the app ‘Pocket Spine Doc’. On different menu pages, the user can access valuable information regarding his back pain, but also fill out questionnaires leading to a personalized exercise program. Furthermore, the user has the opportunity to buy a preventive package with exercises (minor in-app purchasing).

6.0 Apps including Medication Monitoring

Figure 23. Examples of Medication Monitoring of ‘Postapp’

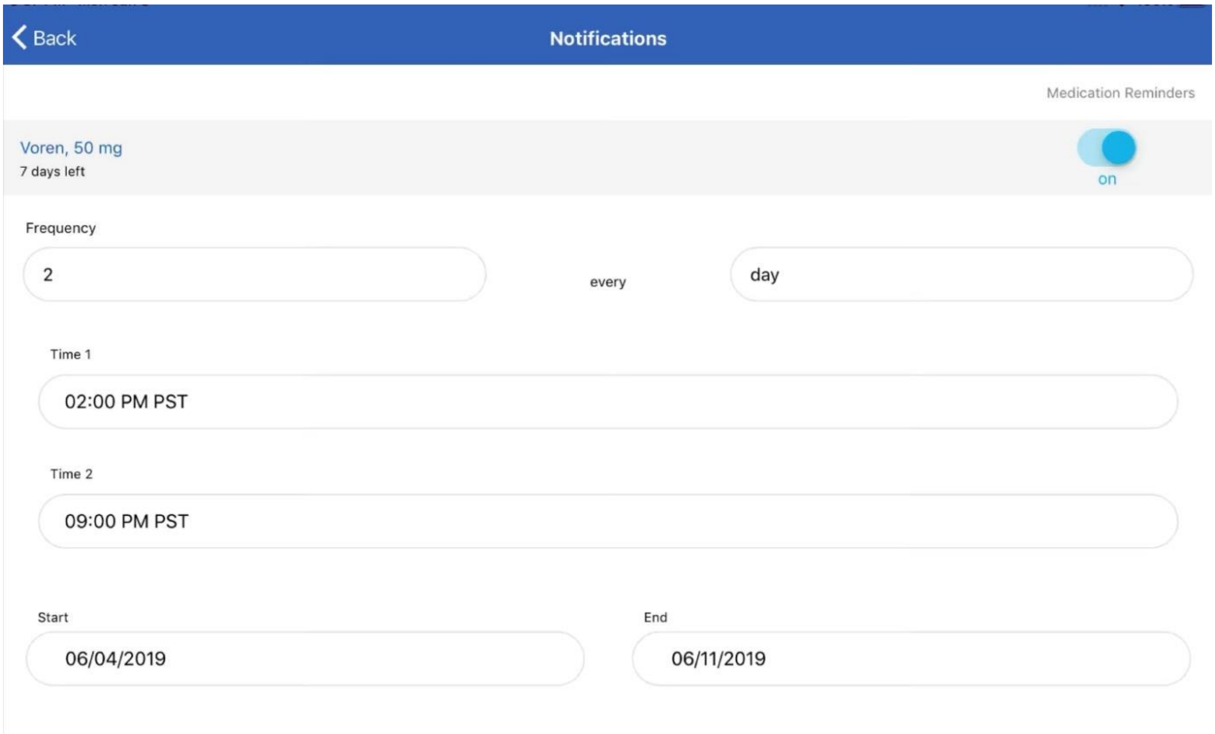

Figure 23. The users of the ‘Postapp’ have the opportunity to monitor their medication, entering the name, dosis, frequency, and time frame and period of intake. Furthermore, the app can send reminders to the user.

Appendix F

App duplicates.

Figure 24. Comparison of Applications Duplicates Part 1

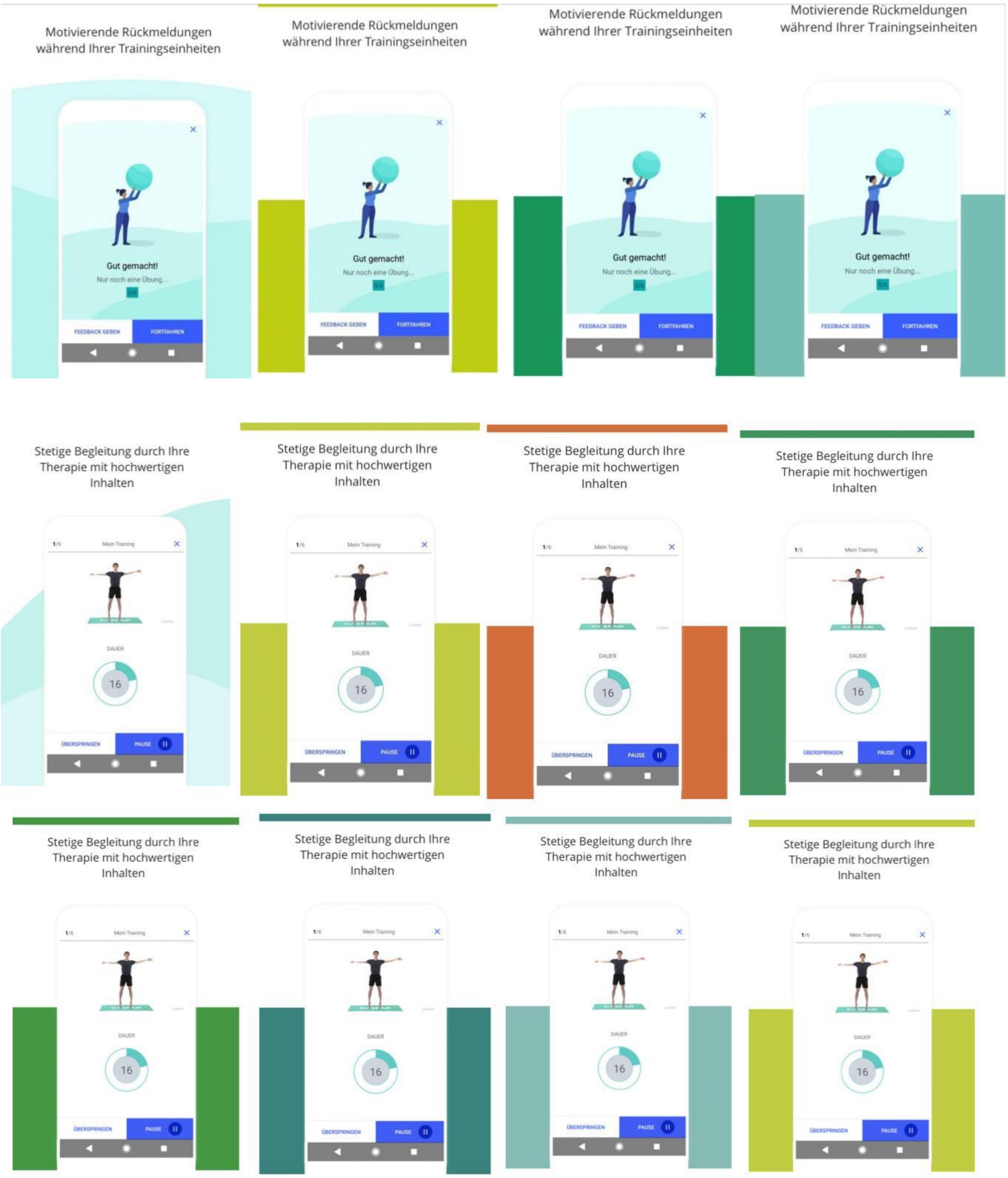

Figure 24. Examples of the Caspar-Health duplicates. Screenshots extracted from the Google Play Store.

Figure 25. Comparison of Application Duplicates Part 2

Figure 25. Examples of all 8 apps and their main home screen.

During the search, 8 duplicates were found, all published by the same company: GOREHA GmbH. The main developer behind these apps is CASPAR, which also developed its own app besides the others (Caspar-Health). It seems that CASPAR adjusted the other apps taking the example of the already developed app. This was also stated on their official website. The content and design of those apps the same, except for the clinic-specific logo, their number of reviews and downloads in the app store. The duplicates are named ‘ZAR Therapie’, ‘f+p Therapie’, ‘Lichtenau Therapie’, ‘Medicos Therapie’, ‘Michaels Kliniken’, ‘Dr. Ebel Therapie’, and ‘Paracelsus Therapie’.