Effectiveness and cost-effectiveness of a web-based routine assessment with integrated recommendations for action for depression and anxiety (RehaCAT+): protocol for a cluster randomised controlled trial for patients with elevated depressive symptoms in rehabilitation facilities

Johannes Knauer, Yannik Terhorst, Paula Philippi, Selina Kallinger, Sandro Eiler, Reinhold Kilian, Tamara Waldmann, Morten Moshagen, Martina Bader, Harald Baumeister

ABSTRACT

Introduction The integration of a web-based computer-adaptive patient-reported outcome test (CAT) platform with persuasive design optimised features including recommendations for action into routine healthcare could provide a promising way to translate reliable diagnostic results into action. This study aims to evaluate the effectiveness and cost-effectiveness of such a platform for depression and anxiety (RehaCAT+) compared with the standard diagnostic system (RehaCAT) in cardiological and orthopaedic health clinics in routine care.

Methods and analysis A two-arm, pragmatic, cluster-randomised controlled trial will be conducted. Twelve participating rehabilitation clinics in Germany will be randomly assigned to a control (RehaCAT) or experimental group (RehaCAT+) in a 1:1 design. A total sample of 1848 participants will be recruited across all clinics. The primary outcome, depression severity at 12 months follow-up (T3), will be assessed using the CAT Patient-Reported Outcome Measurement Information System Emotional Distress—Depression Item set. Secondary outcomes are depression at discharge (T1) and 6 months follow-up (T2) as well as anxiety, satisfaction with participation in social roles and activities, pain, impairment, fatigue, sleep, health-related quality of life, self-efficacy, physical functioning, alcohol, personality and health economic-specific general quality of life and socioeconomic cost and benefits at T1-3. User behaviour, acceptance, facilitating and hindering factors will be assessed with semistructured qualitative interviews. Additionally, a smart sensing substudy will be conducted, with daily ecological momentary assessments and passive collection of smartphone usage variables. Data analysis will follow the intention-to-treat principle with additional per-protocol analyses. Cost-effectiveness analyses will be conducted from a societal perspective and the perspective of the statutory pension insurance.

Ethics and dissemination The study will be conducted according to the Declaration of Helsinki. The Ethics Committee of Ulm University, has approved the study (on 24 February 2021 ref. 509/20). Written informed consent will be obtained for all participants. Results will be published via peer-reviewed journals.

Trial registration number DRKS00027447

INTRODUCTION

Biopsychosocial healthcare in patients with somatic diseases such as musculoskeletal or cardiovascular diseases is faced with the challenge of designing the initiation and
implementation of medical and psychosocial measures in a needs-based manner. Patient-reported outcome measures (PROMs) could become important means to achieve this goal in somatic healthcare. For instance, they can serve as a basis for screening mental health comorbidities, quality of life and functioning in daily living. Based on these test results, treatment planning, process and outcome measurement and quality assurance can take place.

To promote acceptance and to optimise the quality of psychodiagnostics, there is a demand for an economic, resource-saving assessment that minimises the burden on patients (measurement efficiency) and facilitates evaluation and interpretation of results (usefulness). Besides, a reliable assessment of differences between patients (measurement precision) and changes in the condition of patients during and after treatment (change sensitivity of the assessment) is a necessary prerequisite. Paper-and-pencil testing procedures are most commonly used, but have crucial limitations, such as the limited scope or the test load as well as difficulties in collecting these measures before, during and after the treatment process. A computer-based implementation of PROM assessments is considered as a possible solution for an equally precise and time-efficient, user-friendly and useful assessment. Furthermore, a likewise web-based implementation of such an assessment removes time and location dependencies. Hence, assessments of patients before and after the end of treatment could be implemented more easily in routine care settings.

However, PROMs are often static and non-adaptive to the user’s responses, resulting in limited accuracy, presentation of inappropriate items for the individual, and an overall long assessment duration. Computer-adaptive testing (CAT), which is based on item response theory models, is a promising option in this context to substantially reduce the burden on patients (personalised testing) and healthcare institutions (eg, immediate test evaluations). In CAT, the items providing the maximum information about the respective patient are selected and assessed during test administration based on the previous answers of a patient. In this way, besides a considerable reduction in test duration, an estimation accuracy that is equally good or sometimes even better compared with non-adaptive procedures can be achieved.

Whereas PROMs are often used in clinical research, the usefulness and effectiveness of routine PROM assessment in somatic healthcare is still controversially discussed. On the one hand, routine PROMs allow for real-time monitoring. Moreover, PROMs have been shown to lead to improved communication between clinician and patient, decision-making and patient satisfaction with care, and improved health outcome and detection of symptoms and mental comorbidities. On the other hand, there are barriers such as limited ability of clinicians for score interpretation, unfamiliarity with PROM software usage and less time for controlling. Additionally one of the central challenges is the implementation of further evidence-based measures on the basis of the assessment results. Action plans directly derivable from assessment results are regarded as a prerequisite to implement PROMs beneficially. However, there is still an insufficient linkage between assessment results and implementation of existing evidence-based guidelines and recommendations for action in everyday clinical practice.

One way to promote the desired probability of action following diagnostic results could be persuasive design components, such as reminder features that are automatically triggered depending on the test environment. Persuasive designed technological approaches are defined as interactive systems that purposefully influence the user, aiming to change behaviour or attitudes. The provision of computer-based databases and concrete decision-making aids and recommendations for action is seen as one way of reducing these existing barriers. In this context, it could be useful to link the individual test results with therapy standards as well as recommendations for action and guideline knowledge. These have been formulated in particular for the areas of comorbid depression and anxiety in patient populations with somatic diseases. Such a combination could offer the practitioner (1) background knowledge, (2) recommendations for action as well as (3) documentation aids. Ideally, such elements should be directly integrated into testing systems (eg, web based and CAT based) to provide a comprehensive platform from screening to action.

Hence, the aim of the present trial is to examine a persuasive design optimised CAT system (RehaCAT+) providing background knowledge, recommendations for action as well as documentation aids against a standard CAT system (RehaCAT). This will be exemplified with a focus on depression as the primary outcome and anxiety as major mental health comorbidities in cardiological and orthopaedic care. The following research questions will be addressed:

1. Does RehaCAT +improve rehabilitation patients’ depression after 1 year (T3)?
2. Does RehaCAT +improve depression, anxiety, satisfaction with participation in social roles and activities, pain impairment, fatigue, sleep, health-related quality of life, self-efficacy, physical function, and alcohol use at discharge (T1) and 6 months follow-up (T2) as well as 1 year later regarding all the secondary outcomes (T3)?
3. Does RehaCAT +lead to improved documentation and improved follow-up and postrehabilitation recommendations?
4. Does RehaCAT +lead to improved utilisation of rehabilitation therapy standard and guideline compliant healthcare services during and after rehabilitation?
5. What is the cost-effectiveness of RehaCAT +compared with to RehaCAT?
6. What is the acceptance and feasibility of RehaCAT?
7. What are facilitators, hindering factors, mediators and potential risks associated with RehaCAT+?
A web-based CAT system provides a powerful way to assess PROM, however, it is still subject to limitations: (1) it requires active input of the patient—even if reduced through CAT, (2) the assessment is limited to fixed time points, which may lead to long unassessed time intervals in which significant symptom change may occur and (3) due to the nature of self-report the answers by patients may be biased (eg, social desirability or recall bias).\textsuperscript{38-61}

One solution to this could be the addition of ecological momentary assessment and smart sensing to allow for digital phenotyping.\textsuperscript{62,63} Digital phenotyping is defined as the moment-by-moment quantification of the individual health in situ through digital variables and data generated by personal devices (eg, smartphone or smartwatch).\textsuperscript{62,63} First studies show promising results highlighting the potential of this method to complement PROM assessments for monitoring and predicting symptoms with minimal added patient burden.\textsuperscript{64-70} In future the combination of high quality PROM at fixed timepoints combined with continuous monitoring through smart sensing and information from the clinical information system could become a promising data base, which could be used to (1) predict symptom trajectories, (2) build early-detection of adverse events systems (RED-flag) or (3) personalised treatment recommendation systems.\textsuperscript{71-74}

Hence, this study additionally investigates the extent to which smart sensing is suitable for assessing mental health in a routine care setting. In the context of this exploratory study, we will focus on the following research questions:

8. What are the associations between digital markers and health-related variables?
9. Are digital markers suitable for predicting health-related variables and disease or disorder status?
10. What is patient acceptance, adherence and perceived usefulness of smart sensing?

**METHODS AND ANALYSIS**

**Study design**

A two-arm, pragmatic, cluster-randomised controlled trial (cRCT) will be conducted, comparing the experimental group receiving an enhanced version of a PROM system called ‘RehaCAT+’ to the control group receiving the basic version of the PROM system called ‘RehaCAT’ in a 1:1 design (figure 1). See below for detailed description of the experimental and control group.

This cRCT has been approved by the ethics committee of Ulm University (509/20-FSt/Sta) and will be reported in accordance with the Consolidated Standards of Reporting Trials Statement 2010 and the extensions for reporting pragmatic trials and cluster randomised trials.\textsuperscript{75-77} Cost-effectiveness analyses will be reported following the Consolidated Health Economic Evaluation Reporting Standards statement\textsuperscript{78} and the guidelines from the International Society for Pharmacoeconomics and Outcomes Research.\textsuperscript{79} This trial protocol was created according to Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines.\textsuperscript{80} The expected timeline for trial completion is September 2024 with first patient enrolment in July 2022.

**Procedure and recruitment**

The testing systems RehaCAT (control group) and RehaCAT+ (experimental group) will be implemented in the clinic routine of 12 clinics offering cardiological or orthopaedic rehabilitation treatment in Germany. Included clinics pursue medical and occupational oriented stationary rehabilitation according to German ICF diagnosis-based rehabilitation guidelines.\textsuperscript{81} With a psychosocial approach rehabilitation is focused on patients’ impairments (eg, body functions and structure), restoration of activities and removing restrictions of participation.\textsuperscript{82} Accordingly, the treatment in clinics often contains diagnostics, pharmacotherapy, physiotherapy and psychotherapy. Standard stationary stay usually lasts for 3 weeks. The treatment as well as the duration of treatment is expected to vary across patients and between clinics. Treatment will be further described post hoc using the results from the cost-effectiveness questionnaires (see Assessments). Neither the control condition nor the experimental condition will interfere with clinical treatment (see Conditions). For the study, one of two versions of a web-based computer-adaptive diagnostic platform will be implemented within the clinics (see Conditions). Clinical personnel will be trained in an on-site workshop during the implementation phase. The training will cover technical functions of the platform (eg, how new patients can be registered, how patients’ results can be received) as well as recommendations and guidelines for clinical practice.
Open access

(eg, how results should be interpreted, information about national treatment guidelines for mental health). Lastly, clinicians will also be trained in the communication with patients and procedures for patients. After the training, written manuals providing a summary of the workshop will be available in the system for the clinic personnel. Qualification level of clinic personnel operating the system will vary across clinics (eg, nurses, medical doctors, clinical psychologists). This will be monitored and reported (see Usage behavior, acceptance, facilitating and hindering factors). Furthermore, the technical administrator has direct contact options (eg, email) to the research team. The platform is designed so patients can go through the respective version of the testing system to deliver patient-reported outcomes at various points in time. A subset of patients in routine care fulfilling the inclusion criteria (see Inclusion and exclusion criteria) will be included in this study. Study participants will receive all questionnaires from routine care and additional research questionnaires. Routine patients will go through the diagnostic measures at admission (T0) and before discharge (T1), as well as at 6 months (T2) follow-up as part of their clinical routine. Study participants will additionally be assessed at 12 months (T3) follow-up.

Data collection will be digital. Due to the web-based character of the platform, inpatient and outpatient assessments are possible. Clinics are free to implement the admission and discharge assessments as inpatient or outpatient assessments. Data for follow-up will be assessed solely in an outpatient setting. Assessment procedures (eg, inpatient or outpatient assessment at admission) are expected to vary across clinics and will be further described post-hoc. For an explanatory illustration of the assessment procedures, see figure 2.

Inclusion and exclusion criteria

Since this is a cRCT, inclusion and exclusion are differentiated at two levels a) cluster level (ie, eligibility of clinics) and b) patient level: To be eligible, rehabilitation clinics must be located in Germany, provide cardiological or orthopaedic rehabilitation and sign a cooperation agreement with Ulm University. There are no further exclusion criteria for clinics. Within each cluster (ie, rehabilitation clinic) patients who exhibit elevated depression scores (Patient-Reported Outcome Measurement Information System, PROMIS Emotional Distress Depression: T-value ≥65.2) at the initial assessment will be informed about the study and consecutively asked for their participation consent (online supplemental material: SPIRIT Supplement Informed Consent). To be eligible, patients with elevated depression scores must (1) be 18 years or older, (2) have sufficient German language skills, (3) provide an email address, (4) agree to the data privacy and processing procedures according to the European General Data Protection Regulation and (5) sign the informed consent. There are no further exclusion criteria for patients.

Randomisation, allocation and masking

Randomisation and allocation regarding the control (RehaCAT) and experimental group (RehaCAT+) of the 12 participating rehabilitation clinics will be performed by an independent researcher to avoid selection bias. Randomisation will be done on cluster level.

Researchers responsible for randomisation will be obscured to the rehabilitation clinic names and agencies. Randomisation will be done using an automatically created randomisation list.

Figure 2 Procedure.
For the outcome analyses, the conducting analyst will be obscured to group allocation. Patients will remain obscured to their study arm assignment. Neither the clinics (clinic personnel) nor the research team will be obscured to assigned study condition.

Conditions

RehaCAT-Control Group: RehaCAT is a server-based and web-based, device-independent test system, which allows the use of classical test procedures as well as computer-adaptive procedures. RehaCAT does not interfere with the (clinic specific) standard treatment and patients have unrestricted access to treatment as usual. RehaCAT is divided into four user areas: (1) patient, (2) staff, (3) administrator, (4) researcher. The platform allows system administrators to upload and manage patients. Patients go through the diagnostic measures. Clinicians can view the test results of their patients immediately after completion of each assessment point (T0, T1 and T2). Test results consist of a traffic light feedback (green=normal severity, yellow=elevated severity based on clinical cut-off values, red=high severity 2.5 SDs above mean15 83), patients’ test results expressed in T-values combined with clinical cut-off values, and a line graph visualising the results and change over assessment times. For a full overview of the assessment see 2.5.1 and 2.5.2.

RehaCAT+Experimental Group: In addition to the structure and features of RehaCAT, RehaCAT + follows a persuasive design optimised technology (eg, motivation, ability and automatic trigger considering test environment)47 48 to increase the desired probability of action. RehaCAT + offers additional1 system features (automated email reminders for patients),2 clinician features (stored recommendations for action for depression and anxiety based on respective patient results, call to action plans, individualised documentation aids and supporting information material for depression and anxiety)3 and patient features (individual symptomatic information at discharge and T2/3, possible points of contact/help).

The urgency of the recommendation for action (ie, need for in-depth psychodiagnostics) varies depending on screening severity. Additionally, material on handling of psychological burden can be accessed. The material is based on: (1) the rehabilitation therapy standards and framework concepts,84–87 (2) the practice recommendations for orthopaedic and cardiological rehabilitation,54 55 (3) the recommendations for psychodiagnostics in somatic rehabilitation53 and (4) the national S3 guidelines for depression56 and anxiety.57 A summary of the two conditions is provided in figure 3.

The clinics will be compensated with €100 per recruited patient for the resulting hospital expenses in the context of participant recruitment, data collection, study documentation as well as provision of the discharge reports. Study patients will receive an expense allowance of €20 each for their participation in the T2 and T3 measurements.

RehaCAT+ is developed as an open-source platform. It is currently in the certification process according to the medical device regulation (MDR). The platform is developed according to the requirements of the German Medical Devices Act and the MDR. Hence, the software development and validation process is taking the IEC 62304 (safety class B), the GAMP5 (category 4), the general principles of software validation of the Food and Drug Administration (FDA) as well as the Pharmaceutical Inspection Cooperation Scheme 11–3 into account. Furthermore, technical requirements and standards for the interoperability between different medical devices (eg, HL7 FHIR) are under development. The certification process of the platform is planned to be completed in 2022.

Sample size and study power

The sample size calculation is based on the primary outcome, depression severity 12 months after the end of rehabilitation. It is assumed that the experimental group is superior to the control group in the way that patients will show fewer depressive symptoms at 12 months follow-up. In view of the additive study design testing for incremental benefit of RehaCAT + over RehaCAT and the distal outcome, a small additional effect of d=0.24 compared with the standard condition is regarded as clinically significant following the recommendation of Cuijpers et al.88 With 2×6 cluster-randomised rehabilitation clinics, each clinic requires a sample of 110 (SD=25) participating rehabilatants with elevated depression scores to achieve a test power of 80% given an alpha error (two-sided) of 0.05, an estimated ICC of 0.02 98 99 and an assumed correlation with baseline depression scores of 0.50. With an estimated drop-out rate (rehabilitation
start-end) of 20%\textsuperscript{91} and the assumption of a doubling drop-out rate by T3, a total sample of N=1848 rehabili-
tants is required.

**Assessments**
Quantitative outcome assessment will be performed at baseline/beginning of rehabilitation (T0), at discharge/ end of rehabilitation (T1), and at 6 and 12 months (T2, T3). Primary outcome is the depression severity score twelve months after rehabilitation (T3). An overview is provided in table 1.

**Primary outcome**
Depression severity will be assessed with the computer-adaptive PROMIS Emotional Distress-Depression Item Set including an item bank with 28 items that capture negative mood, decrease in positive emotions, cognitive deficits, as well as negative self-image and negative social cognition.\textsuperscript{92} All items are rated on a five-point response scale asking respondents to rate the frequency of their symptoms in the past 7 days (never, rarely, often, sometimes, always). A Cronbach’s alpha of 0.99 was found for the internal consistency of the item set.\textsuperscript{93}

**Secondary outcomes**
Anxiety will be assessed with the computer-adaptive PROMIS Emotional Distress-Anxiety Item Set including 29 items that assess emotional distress caused by hypersensitivity, anxiety and stress, as well as associated somatic symptoms.\textsuperscript{92} All items are rated on a five-point response scale asking respondents to rate the frequency of their symptoms in the past 7 days (never, rarely, often, sometimes, always). The internal consistency of the item set was found to be good.\textsuperscript{92}

Pain impairment will be assessed with the computer-adaptive PROMIS Pain Interference Item Set including 40 items that capture self-assessment of the consequences of pain in one’s life. This includes the extent to which pain interferes with engagement in social, cognitive, emotional, physical, as well as leisure activities.\textsuperscript{95} The items refer to the past 7 days and are rated on three different five-point Likert scales. The internal consistency of the item set was found to be good.\textsuperscript{95}

Fatigue will be assessed with the computer-adaptive PROMIS Fatigue Scale comprising 95 items and measures both fatigue experience and the impact of fatigue on daily life and functionality. The intensity, frequency and duration of fatigue were graded on a five-point response scale. Reliability was estimated to be \(\alpha > 0.90\).\textsuperscript{96}

<table>
<thead>
<tr>
<th>Variable</th>
<th>Instrument</th>
<th>CAT</th>
<th>Time of measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>PROMIS emotional distress—depression</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Anxiety</td>
<td>PROMIS emotional distress—anxiety</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Satisfaction with participation in social roles and activities</td>
<td>PROMIS satisfaction with social roles and activities</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Pain impairment</td>
<td>PROMIS pain interference</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Fatigue</td>
<td>PROMIS fatigue</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sleep</td>
<td>PROMIS sleep disturbance</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Health-related quality of life</td>
<td>PROMIS global health</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>PROMIS self-efficacy general</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Physical Function</td>
<td>PROMIS physical function</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Alcohol use</td>
<td>AUDIT-10</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Personality</td>
<td>BFI-10</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Generic quality of life</td>
<td>EQ5D-5L</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Health and social services use and costs</td>
<td>CSSRI</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Medical record data</td>
<td>Provided by clinicians (eg, discharge reports)</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

PROMIS, Patient Reported Outcome Measurement Information System; CAT, Computer-Adaptive Patient Reported Outcome Test; CSSRI, Client Sociodemographic and Service Receipt Inventory; EQ5D-5L, European Quality of Life 5 Dimension - 5 Level Questionnaire; PROMIS, Patient Reported Outcome Measurement Information System.
Sleep will be assessed with the computer-adaptive PROMIS Sleep Dimension that measures subjective sleep quality and quantity and sleep-related impairment in daily functioning. The scale comprises 27 items and is rated on a response scale from 1 (not at all) to 5 (very much or always). It is a validated instrument and has good psychometric properties with $\alpha=0.90$.97 98

Health-related quality of life will be assessed with the PROMIS scale on global health aspects (Global Health) is used. The scale includes 10 items that capture global physical health (physical health, physical functioning, fatigue, pain), and global mental health (general quality of life, mental health, satisfaction with social activities and relationships, and emotional distress).99 Nine items are scored on a response scale of 1–5, and the item assessing pain is scored from 0 to 10. Internal consistency was estimated to be good with $\alpha=0.82$.99

Self-efficacy will be assessed with the short scale of the PROMIS General Self-Efficacy Scale which contains four items. It can be used to assess how much confidence one has in one’s own abilities to master tasks and situations. Items are rated on a response scale from 1 (I am not at all confident) to 5 (I am very confident). Internal consistancy was estimated to be high ($\alpha=0.96$).100 101

Physical function will be assessed using the computer-adaptive PROMIS Physical Function Item Set that measures the ability to perform daily life activities that require physical activity such as walking or lifting objects. It comprises 164 items that are assessed on a 5-point scale ranging from 1 (Without any difficulty) to 5 (Unable to do). Internal consistency was found to be very good with $\alpha=0.88$.102

Alcohol use will be assessed with the Alcohol Use Disorders Identification Test which is a 10-item questionnaire developed by the WHO as a screening tool for hazardous and harmful alcohol use. Responses to each question are scored from 0 to 4, allowing a maximum of 40 points. Reliability has been investigated in some studies and is considered good with a median of $\alpha=0.80$.103

Discharge reports will be analysed regarding (1) frequency of documented screening results, (2) therapy standard and guideline appropriate therapeutic services (documented services/therapy standard recommendations as a function of depression and anxiety results), (3) therapy standard and guideline appropriate follow-up and postrehabilitation recommendations (documented recommendations/therapy standard/guideline recommendations as a function of depression and anxiety results).

**Moderators**

As potential moderators, sociodemographic data (age, gender, nationality) and personality will be recorded at admission (T0). Personality will be assessed with the BFI-10 (Big Five Inventory-10), a short version of the BFI that has good psychometric properties and a retest-reliability of $\alpha=0.73$.104 Additionally, data from medical records will be used as moderators (eg, indication area orthopaedic or cardiology, chronic conditions, rehabilitation duration).

**Health economics**

Generic quality of life will be assessed with the European Quality of Life 5 Dimension-5 Level Questionnaire from the EuroQol foundation (www.euroqol.org).105 The five dimensions surveyed are mobility, self-care, general activities, pain/physical discomfort and anxiety/dejection. Each of the five areas is ranked on a five-point scale. Based on the answers, the respective health status is recorded.106

Health and social services use and costs will be assessed with the Client Sociodemographic and Service Receipt Inventory which is a standardised but adaptable inventory. Five domains are queried, including sociodemographic information, usual living situation, income and employment status, use of mental health services, and medication treatment.107

**Usage behaviour, acceptance, facilitating and hindering factors**

Questions about usage behaviour, potential risks of the platforms, as well as barriers and facilitators to implementation, will be elicited based on qualitative semistructured interviews conducted with both patients and clinic staff centrally involved in the implementation of RehaCAT and RehaCAT+. The semistructured interviews will be conducted with the help of an interview guide based on existing instruments of previous studies.55 108

**Smart sensing substudy**

Smart sensing data will be collected actively and passively via the mobile application AWARE for smartphones. The AWARE Framework is an open-source framework that allows to passively collect digital behavioural data via an app as well as to send ecological momentary assessments (eg, short questions: “how are you feeling right now?”) to the app user for answering.109 The AWARE framework has been tested in previous studies94 106 109 110 without technical, privacy or ethical issues. All collected data will be stored pseudonymised and personal data (eg, contact numbers) will be anonymised using the secure hash algorithm 256 (SHA-256).109

After completing the diagnostic measures at T0, T1 and T2, all patients will be informed in the RehaCAT (+) system about the optional mobile sensing substudy. If interested, they can provide an email address to receive further information on the study and a study invitation. This is independent from study participation in the cRCT. Therefore, both routine care patients and patients partaking in the cRCT will be able to participate. Participants who provide their informed consent will be instructed to install the research application on their personal smartphones. After installation participants will be able to choose which data points will be collected over the next 6 months.

**Active assessment**

Gender, age and personality with the BFI-10 will be assessed104 once after installing the application.
Furthermore, acceptance of and satisfaction with smart sensing will be measured using the Unified Theory of Acceptance and Use of Technology questionnaire,\textsuperscript{111,112} satisfaction with the research application will be measured with the User Version of the Mobile Application Rating Scale.\textsuperscript{113} Both questionnaires will be assessed once after 6 months before deinstalling the application.

The following clinical questionnaires will be assessed every 2 weeks: Depression (dimensional and categorical) with the PHQ-8 (if PHQ-2 score >2), anxiety with the GAD-7,\textsuperscript{114–116} stress with the PSS-10,\textsuperscript{117} sleep with the ISI,\textsuperscript{118} loneliness with the UCLA three-item version.\textsuperscript{119}

Every morning, participants will be asked short questions about mood (valence), drive (arousal), control, unpredictability, stress and sleep, at midday about mood (valence), drive (arousal), control, unpredictability and stress, and in the evening, participants are again asked about mood (valence), drive (arousal), control, unpredictability, stress and activity during the day. This assessment is based on previous studies.\textsuperscript{65,68,121,122}

**Passive outcomes**

The research app allows to track a broad range of sensors (accelerometer, application usage, barometer, battery, Bluetooth, communication, gravity, gyroscope, light, locations, magnetometer, network, proximity, rotation, screen sensor). However, each user will be able to freely decide which sensors are activated and access permissions can always be activated and deactivated without giving reasons. In addition, sensible location data (eg, Global Positioning System (GPS) coordinates) will be obscured, so pseudonymisation can be upheld all the time.

The following digital markers can be collected (depending on permissions of user): frequency and duration of smartphone and individual app usage, frequency and duration/length of calls and text messages, randomly distorted GPS, and type of movement.

**Data management and data sharing plan**

Data collection will be completed online using the server-based system RehaCAT(+) and the research application in pseudonymised form. Retrieved data will be stored encrypted by responsible employees. All data will be anonymised after completion of the trial. Furthermore, an independent data safety and monitoring board (DSMB) with longstanding experience in clinical trials has been established. The function of the DSMB is to monitor the course of the study and, if necessary, give recommendations to the steering committee for discontinuation, modification or continuation of the study.

Individual participant data will be made available on request after de-identification beginning 12 months following article publication of the effectiveness paper. Data will be made available to researchers who provide a methodologically sound proposal, not already covered by others. Proposals should be directed to HB. Data requestors will need to sign a data access agreement.

Provision of data is subject to data security regulations. Investigator support depends on available resources.

**Measures to reduce methodological sources of error**

Selection bias: Randomisation and allocation regarding the group allocation (RehaCAT / RehaCAT(+)) of the participating rehabilitation clinics will be done by an independent researcher. Performance bias: Rehabilitation staff centrally involved in the implementation will be trained along training materials, as well as continuously supervised regarding the training materials. RehaCAT(+) and its application will be described in detail in a test manual. Deviations from the test manual will be recorded and formatively reduced during the implementation process of RehaCAT(+) in the individual clinics. Contamination bias: Cluster randomisation is used to avoid study arm contamination. Detection bias: rating procedures (analysis of discharge reports) are performed by independent raters who are obscure to the study arm affiliation of the clinics and thus the patients. Patients also remain obscured to their study arm assignment, as the differentiation between RehaCAT and RehaCAT(+) is not obvious to them. Obscuring will not be realised for clinic personnel only. Reporting bias: A detailed definition of all methodological aspects of the present clinical study is provided in this study protocol, submitted for publication prior to randomisation start. Evaluating representativeness: To assess the representativeness of the results, quantitative and qualitative analyses will be performed regarding the reasons that may lead to a non-existent 100% exhaustion in the routine assessment.

**Statistical analyses**

**Clinical analyses**

Study data will be centrally processed and analysed by an independent researcher. Missing values and missingness patterns will be explored and analysis will be adjusted accordingly using multiple imputation strategies (based on heteroscedastic two-level linear models considering the metric of outcome). The analysis will follow the intention-to-treat principle. In addition, per-protocol analyses will be conducted. The primary outcome as well as all other continuous outcomes will be analysed based on Hierarchical Linear Models considering cluster structure and baseline values. Binary outcomes will be analysed using mixed logistic regression models. Moderator and mediator analyses will be performed to determine differential effects with respect to key sociodemographic and medical variables.

The effect of study participation will also be measured (participation rate at T1-3) in order to be able to make statements about the transferability of the results from the present randomised study to routine care without research support.

**Health economic evaluation**

In the health economic evaluation, an incremental cost–utility analysis will be performed from the societal...
perspective, as well as from the perspective of the German statutory pension insurance (SPI) according to the net benefit approach. The necessary maximum willingness to pay (MWTP) for a clinical improvement of depressive symptoms by 50% (=response) and for the gain of a quality-adjusted life-year will be determined. The estimation of the stochastic uncertainty will be done by means of nonparametric bootstrapping. The interpretation of the results is based on cost-effectiveness acceptance curves. These provide information on how high the MWTP must be to be judged cost-effective with a probability of 95%, or with what probability a pre-determined MWTP is judged to be cost-effective. Following international guidelines, a value range of the MWTP between €0 and €1,250,000 is chosen. The analysis of the health economic relevance of moderator and mediator variables will be performed by means of net benefit regression models for net benefit ratios between €0 and €1,250,000. The analysis from the macroeconomic perspective will consider all direct and indirect disease costs, the analysis from the perspective of the SPI will take the disease costs to be borne by the SPI (eg, for medical rehabilitation services) as well as the costs for the testing platform to be borne by the SPI into account.

Qualitative data analysis

Qualitative interviews of patients and clinic staff will be conducted and analysed. The analysis of qualitative data will be based on qualitative content analysis. An inductive-deductive approach will be applied along the theory-based interview guide. Reliability of results will be established (indicated by intercoder agreement) with two independent raters coding all transcripts on the basis of coding guide and rules. This coding guide will be developed in an iterative process with consensus finding.

Smart sensing

Identification of prognostic factors

Correlation analysis and multilevel regression models will be used to identify bivariate relationships between PROMs, digital markers, disease-free survival and postoperative complications. The presence of missing values and missingness patterns will be investigated and analysis will be adjusted accordingly (eg, multiple imputation).

Prediction modelling

Significance test-based methods as well as machine learning models will be used. In case of nested data structure (eg, development of quality of life over time with repeated measures) multilevel regression models will also be used for prediction. For continuous outcomes (eg, depression severity) linear models will be used, while logistic models will be applied for dichotomous outcomes (eg, depression state yes or no).

In addition to these significance test-based methods, machine learning models will be used. Machine learning is an iterative process and several modelling approaches will be tested (eg, K-Nearest Neighbour algorithm or gradient-boosted trees). However, since the field is rapidly developing, we cannot a priori define the exact approaches that will be used. Hyperparameter optimisation will be conducted using grid-search.

Patient and public involvement

Patient and public involvement (PPI) representatives have provided input to this study in several stages. Results of previous projects including patient feedback, were used to further develop and optimise study design and procedures. PPI representatives (eg, as members of an advisory board) are included to improve usability, design and comprehensibility but have no influence on the outcomes, data analysis methods or study design.

ETHICS AND DISSEMINATION

Ethical approval has been obtained from Ulm University on 24 February 2021 (ref. 509/20). The study is conducted according to the Declaration of Helsinki. Informed consent will be obtained from all participants.

Results will be published in peer-reviewed journals. They will also be made known through local conferences and research seminars, national and international scientific congresses, and through direct and indirect contacts with clinicians, public health managers and other healthcare professionals.

Author affiliations

1 Department of Clinical Psychology and Psychotherapy, Ulm University, Ulm, Germany
2 Department of Psychiatry and Psychotherapy II, Ulm University, Ulm, Germany
3 Department of Psychological Research Methods, Ulm University, Ulm, Germany

Acknowledgements

Authors would like to thank the project members Carolin von Gottberg, Minh Duc Duong, Erce Rodopman, Tobias Sterns, Aydin Spieler, Rebecca Knoll and Robin Kraft for their technical support and development of RehaCAT.

Contributors

HB is principle investigator of RehaCAT+, HB, RK and MM obtained funding for this study. HB, JK, YT, PP, SK, MM, MB, RK and TW contributed to the study design. HB, SE, JK, YT, PP and SK developed the platform RehaCAT+. MM and RK contributed to the design of the effectiveness and health economic evaluation. JK drafted the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Federal Ministry of Education and Research (BMBF, 01GX1901).

Disclaimer

BMBF had no role in study design, decision to publish or preparation of this manuscript. BMBF will not be involved in data collection, analyses, decision to publish or preparation of future papers regarding this study.

Competing interests

Authors of the manuscript were partly involved in the development of RehaCAT+. HB has been the beneficiary of study support (third party funding) from several public funding organisations in the context of research on computer-adaptive testing and patient-reported outcome systems.

Patient and public involvement

Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication

Consent obtained directly from patient(s).

Provenance and peer review

Not commissioned; externally peer reviewed.

Supplemental material

This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content...
includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access

This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Johannes Knauer http://orcid.org/0000-0002-7679-7090
Yannik Terhorst http://orcid.org/0000-0003-4091-5048
Morten Moshagen http://orcid.org/0000-0002-2929-7288

REFERENCES


8 Field J, Holmes MM, Newell D. Proms data: can it be used to make decisions for individual patients? A narrative review. Patient Relat Outcome Meas 2019;10:233–41.


10 Gamper EM, Martini C, Petersen MA. Do patients consider computer-adaptive measures more appropriate than static questionnaires? J Patient Reported Outcomes 2019;3.


42 Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a


DGPN, BAK, KWB, AWMF, AkbA, BPKD DAGSHG, DEGAM, DGPM, DGRW, S3-Leitlinie/Nationale VersorgungsLeitlinie Unipolare depression. Langfassung 2015.


11


126 Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. In: lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics, 1995.

---

BMJ: first published as 10.1136/bmjopen-2022-061259 on 23 June 2022. Downloaded from http://bmjopen.bmj.com/ by guest. Protected by copyright.