


BMJ Open STEP.De study – a multicentre cluster-randomised effectiveness trial of exercise therapy for patients with depressive symptoms in healthcare services: study protocol

Andreas Heissel ¹, Anou Pietrek,¹ Melanie Schwefel,¹ Kahar Abula,¹ Gregor Wilbertz,² Stephan Heinzl,² Michael Rapp¹

To cite: Heissel A, Pietrek A, Schwefel M, *et al.* STEP.De study—a multicentre cluster-randomised effectiveness trial of exercise therapy for patients with depressive symptoms in healthcare services: study protocol. *BMJ Open* 2020;**10**:e036287. doi:10.1136/bmjopen-2019-036287

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-036287>).

Received 09 December 2019
Revised 17 February 2020
Accepted 12 March 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Social and Preventive Medicine, Department Sport and Health Sciences, Faculty of Human Sciences, University of Potsdam, Potsdam, Brandenburg, Germany

²Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany

Correspondence to

Dr Andreas Heissel;
andreas.heissel@uni-potsdam.de

ABSTRACT

Introduction Although exercise therapy has widely been shown to be an efficacious treatment modality for depression, evidence for its effectiveness and cost efficiency is lacking. The Sport/Exercise Therapy for Depression study is a multicentre cluster-randomised effectiveness trial that aims to compare the effectiveness and cost efficiency of exercise therapy and psychotherapy as antidepressant treatment.

Methods and analysis 480 patients (aged 18–65) with an International Classification of Diseases diagnosis associated with depressive symptoms are recruited. Up to 30 clusters (psychotherapists) are randomly assigned to allocate patients to either an exercise or a psychotherapy treatment as usual in a 2:1 ratio. The primary outcome (depressive symptoms) and the secondary outcomes (work and social adjustment, quality of life) will be assessed at six measurement time points (t0: baseline, t1: 8 weeks after treatment initiation, t2: 16 weeks after treatment initiation, t3/4/5: 2, 6, 12 months after treatment). Linear regression analyses will be used for the primary endpoint data analysis. For the secondary endpoints, mixed linear and logistic regression models with fixed and random factors will be added. For the cost efficiency analysis, expenditures in the 12 months before and after the intervention and the outcome difference will be compared between groups in a multilevel model. Recruitment start date was 1 July 2018 and the planned recruitment end date is 31 December 2020.

Ethics and dissemination The study protocol was approved by the ethics committee of the University of Potsdam (No. 17/2018) and the Freie Universität Berlin (No. 206/2018) and registered in the ISRCTN registry. Informed written consent will be obtained from all participants. The study will be reported in accordance with the Consolidated Standards of Reporting Trials and the Recommendations for Interventional Trials statements. The results will be published in peer-reviewed academic journals and disseminated to the public.

Trial registration number ISRCTN28972230.

Strengths and limitations of this study

- To our knowledge, this is the first study to examine the effectiveness and cost efficiency of exercise therapy for patients with depressive symptoms and its implementation into healthcare services as a low-threshold, quickly accessible treatment option.
- Implementation of a pragmatic trial design ensures that the methodological features of the trial are as closely aligned with routine practice as possible, by following medical and psychological guidelines while still adhering to the methodological requirements of a rigorous randomised controlled trial research design.
- Cost efficiency analysis complements evidence on effectiveness and provides valuable economic insights to guide decisions regarding resource allocation. Anticipated limitations include: open-label, single-blind study design, lack of patient public involvement in the development of the study protocol and exercise regimen, and a comparison which is based on the intervention period (16 weeks).

INTRODUCTION

According to WHO,¹ depression is ranked as the single largest contributor to disability worldwide, affecting over 300 million people and accounting for 7.5% of all years lived with disability.² Beyond the personal suffering, depression is associated with high unemployment and reduced performance at work.³ Among the treatments for major depression disorder (MDD) in medical care, psychotherapy is one of the main therapeutic approaches.⁴ Although its effectiveness has widely been proven and accepted,^{5 6} psychotherapeutic treatment is often associated with long waiting times for patients (3–6 months in Germany)⁷ and substantial costs for the healthcare system.⁸ These challenges in the standard care of MDD illustrate the need for

other potential treatment strategies. In this regard, exercise therapy has been proposed as an useful intervention option for several years, as it is easily accessible,⁹ improves physical health (in a population at higher risk of comorbidities)¹⁰ and does not have the side effects that come with antidepressants, especially in the treatment of mild-to-moderate depressive symptoms.⁹ However, much of the supporting evidence for exercise interventions has come from strictly controlled efficacy trials under ideal conditions (phase III trials), and information on the effectiveness and cost efficiency of exercise therapy under conditions of actual use (phase IV trials) compared with psychotherapy in routine clinical practice is lacking.⁹

Efficacy of exercise as a treatment for MDD

The German National Supply Guideline for Unipolar Depression (S3 Guideline)¹¹ and the European Psychiatric Association¹² recognise exercise therapy as a complementary treatment option for MDD. Several recent meta-analyses have demonstrated moderate-to-large effects of exercise in the treatment of depression.^{13–19} While the antidepressant effect of aerobic exercise has the strongest empirical evidence, recent meta-analyses suggest that strength training is also effective^{18 20} and that the combination of strength and aerobic exercise might result in the largest effects.¹⁵ Other forms of physical exercise, like yoga, that can be classified as mindfulness-based exercise forms, were also shown to be moderately effective in the treatment of depression.²¹ Regarding the underlying mechanisms, there is correlational evidence for an association between exercise, performance in general cognitive domains and specific emotional regulation capacities.²² In fact, it has been shown that physical exercise can improve performance in a subsequent working memory test.²³ Moreover, improved working memory results in better emotional regulation with regards to both psychological and biological markers.^{24 25} Studies that translate these presumed mechanisms to the effect of exercise on depressive symptoms are missing. When it comes to the method of delivery, larger effect sizes and lower dropout rates were found for exercise therapies that were supervised by professionals¹⁴ (ie, exercise physiologists, exercise therapists, physiotherapists, etc). This is consistent with previous findings that compared with home-based trainings, the supervision of exercise treatment is a key predictor of successful treatment outcome in depression trials.²⁶ Regarding a direct comparison between exercise and psychotherapy, empirical evidence is sparse to date with methodological constraints and a comparison with psychotherapy in routine care is missing.

There is several evidence in the literature for the long-term effect of exercise therapy on depression, with pragmatic randomised controlled trials (RCTs) following rigorous methodological guidelines and with the follow-up periods ranging from 3 month, 6 months or even 12 months.^{27–29} However, on the other hand, the remission rate of psychotherapeutic treatments in MDD is

lower than 50%,^{30 31} which inevitably affects its long-term effect.

Limitations of efficacy trials and the need for pragmatic effectiveness trials of exercise treatment for MDD

The current evidence base for efficacy trials is fraught with some methodological limitations that concern several of the participants, interventions, comparisons, outcomes, and study designs (PICOS) criteria,³² which need to be considered in order to clearly understand the impact of exercise on depression (for a critical review).⁹ In particular, most studies have adopted restrictive inclusion and exclusion criteria for participants, with the result that the participants differ substantially from those seen in primary care (ie, self-selection bias).⁹ With regard to comparison groups, most efficacy trials typically compare exercise treatment groups to control groups with different exercise modalities and/or intensities, whereas non-physically active comparison groups are preferred to minimise treatment crossovers (ie, change of patients from the experimental arm to the standard arm after treatment initiation) and to curtail the underestimation of the effects of exercise interventions.⁹ In terms of outcomes, the majority of studies tend to choose outcome measures that do not reflect most of the real-life concerns of patients, clinicians and policy-makers: more pertinent outcome measures such as quality of life, return to work, readmission to hospital, cost, and so on, did not receive the attention they deserve.³³ Taking treatment cost as an example, to date little is known about the practical question of the cost efficiency of exercise therapy. Cost efficiency is an important measurement of programme resource use, complementing evidence on efficacy and providing valuable economic insights to guide decisions regarding resource allocation and priority setting.³⁴ Lastly, concerning study design, pragmatic trials that allow scalable replication and implementation in routine practice are scarce. Ensuring the pragmatic effectiveness of a treatment modality in regular healthcare services depends not only on the efficacy of the treatment, but also on multiple additional factors, such as patient adherence, diagnostic accuracy and provider compliance.³⁵ Another limitation is the short time span in which outcomes are measured and the lack of a sufficient follow-up period after the intervention. Due to the high frequency of relapse and remittance in patients suffering from depressive disorders,³⁶ longer follow-up assessments can provide the needed key information about the lasting effects of exercise intervention programs.

In short, based on the PICOS criteria, the current evidence on efficacy is insufficient to guide physicians and policy-makers in choosing the optimal treatment for patients. Therefore, conducting methodologically rigorous, pragmatic RCTs to investigate the effectiveness of exercise treatment in usual care settings is important. It fills the implementation gap between evidence and practice and provides crucial information about the treatment's generalisability in the healthcare system.

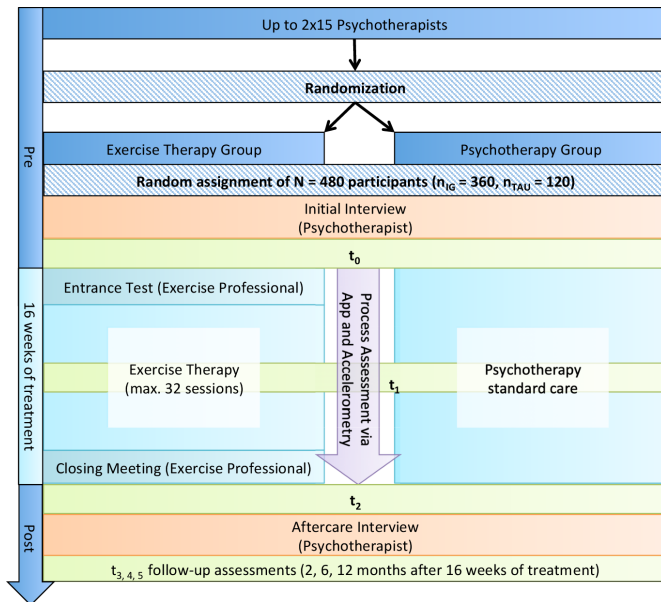


Figure 1 Study design procedure.

Pragmatic effectiveness trial of exercise as a treatment for depressive symptoms: the STEP.De study

The Sport/Exercise Therapy for Depression (STEP.De) study is a multicentre cluster-randomised effectiveness trial. The aim is to evaluate exercise therapy as a treatment strategy in community-based healthcare in comparison to standard-care psychotherapy. Additionally, underlying mechanisms are examined during the treatment period via ecological momentary assessment (EMA). We examine the following hypotheses: (1) after 16 weeks of intervention and at 2-month, 6-month, and 12-month follow-up, both intervention groups (IGs) show significant improvement in their depressive symptoms, the ability to work, health-related quality of life and psychopathological symptoms. (2) After 16 weeks of intervention and at 2-month, 6-month, and 12-month follow-up, exercise therapy is not inferior to psychotherapy (a) in decreasing depressive and other psychopathological symptoms, (b) in increasing patients' ability to work and health-related quality of life, (c) in remission rates and (d) in its economic cost efficiency. Based on EMA, we expect to confirm that (3) physical activity has an immediate positive effect on patients' emotional state, cognitive and emotional regulation capacities, and (4) symptom improvement (over 4 months of treatment) is driven by improvements in emotion regulation, which is itself driven by improvements in cognitive control following physical activity.

METHODS AND ANALYSIS

Study design

In order to ensure the concealment of allocation, a blinded independent researcher will randomly assign up to 30 clusters (psychotherapists), stratified by location, to one of two groups: one group directs their patients to an exercise therapy group (IG) which provides supervised

exercise therapy. The other group of psychotherapists provides individual cognitive behavioural or psychodynamic psychotherapy (control group, TAU). In total, 480 patients will be randomly allocated in a 2:1 ratio to either the exercise treatment or psychotherapy as usual by two groups of psychotherapists for the duration of the study. Both interventions will take place over 16 consecutive weeks and participants in each treatment arm will receive follow-up assessments for 12 months in total. At six measurement time points (t_0 : baseline, t_1 : 8 weeks after treatment initiation, t_2 : 16 weeks after treatment initiation, $t_3/4/5$: follow-up 2, 6, and 12 months after treatment, respectively), primary and secondary outcome measures will be assessed via online assessment and structured telephone interviews. An overview of the study design is shown in figure 1. During the treatment period, patients will participate in a smartphone-based, extensive EMA employing regular evening assessments (every 2–4 days), working memory tests (every 5–9 days), and continuous mobile sensing (see table 1). EMA is administered using the mobile apps *movisensXS*, *presentation* and *StepProcess* (in-house application) installed on the participants' phones if possible (Android) or study phones otherwise (phone ownership will be added as a covariate to the analyses). In addition, during 3 periods of 7 days each (first and last treatment weeks, one random week in between), questions will be sampled at a higher rate (four times a day) and will be accompanied by continuous accelerometry (using *GENEActive* devices).

The study will be an open-label study and the therapists and patients will not be double blinded to the group affiliation. Assessors conducting telephone assessments will be blinded to the group affiliation of the participants. Nevertheless, all involved parties, including patients, psychotherapists, exercise professionals and assessors will be informed and aware of the evidence-based assumption that the two treatment modalities are similar in terms of their treatment effect, so they will expect both treatments to be equally helpful in reducing depressive symptoms.

The study protocol is designed, constructed and reported according to the recommendations given in the Standard Protocol Items: Recommendations for Interventional Trials³⁷ and Consolidated Standards of Reporting Trials (CONSORT)³⁸ statements, especially its adapted extensions for pragmatic trials,³⁹ cluster RCTs (CRCTs)⁴⁰ and non-inferiority trials⁴¹ (see online supplementary tables 1, 2). To ensure the quality of the study, the Pragmatic-Explanatory Continuum Indicator Summary 2 tool⁴² (the present trial gets a score of 5, that is, high pragmatic quality), the Physiotherapy Evidence Database scale⁴³ (the present trial gets a score of 5–8, ie, low risk of bias) and WHO Trial Registration Data Set will be used (see online supplementary tables 3–5). The study is entered in the ISRCTN registry (<https://doi.org/10.1186/ISRCTN28972230>) and the CONSORT extension for journal abstracts is attached (see online supplementary table 6) and will be used (see online supplementary

Table 1 Overview of outcome measures

Measures	Instrument	Abbreviation	Measurement time points					
			t ₀	t ₁	t ₂	t ₃	t ₄	t ₅
Primary outcome								
Depression severity	Beck Depression Inventory ⁵¹	BDI-II	X	X	X	X	X	X
Secondary outcome								
Work and social adjustment	Work and Social Adjustment Scale ⁵³	WSAS	X	X	X	X	X	X
Health-related quality of life	EuroQol Five-Dimensional Questionnaire ⁶²	EQ-5D	X	X	X	X	X	X
Health-related quality of life	12-Item Short Form Survey ⁵⁶	SF-12	X		X		X	
Basic psychological needs frustration and satisfaction	Basic Psychological Need Satisfaction and Frustration Scale ⁵⁴	BPNSFS	X		X		X	
Sports motivation	12-Item Short Form Behavioral Regulation Questionnaire ⁶⁰	BRQ-12	X		X		X	
Mindfulness	Mindfulness Attention Awareness Scale ⁶³	MAAS	X		X		X	
Self-efficacy	General Self-Efficacy Scale ⁵⁸	GSE-6	X		X		X	
Healthcare climate	Health Care Climate Questionnaire ⁵⁹	HCCQ	X	X	X			
Work ability	WHO Disability Assessment Schedule 2.0 ⁴⁷	WHODAS 2.0	X		X		X	
Depressive symptoms	Hamilton Rating Scale for Depression ⁴⁸	HAM-D	X		X		X	
Psychopathological symptoms	Symptom Checklist 90 scale ^{49, *}	SCL-90	X		X		X	
Physical activity	International Physical Activity Questionnaire ⁵⁰	IPAQ	X		X		X	
Process evaluation			Collected through			Measurement time points		
Physical activity	Accelerometry		Continuously during 3 periods of 7 days each (first and last treatment weeks, one random week in-between)					
Emotion	Ecological momentary assessment		At four random time points per day during 3 periods of 7 days each (first and last treatment weeks, one random week in between)					
Emotional regulation	Ecological momentary assessment							
Cognitive control	Ecological momentary assessment							
Depressive symptoms	Ecological momentary assessment		At evening assessment twice per week during the entire intervention period					
Emotional regulation	Ecological momentary assessment							
Physical activity	Ecological momentary assessment							
Cognitive control	Ecological momentary assessment							
Self-efficacy	Ecological momentary assessment							
Therapy motivation	Ecological momentary assessment							
Therapy progress	Ecological momentary assessment							
Adverse effects	Ecological momentary assessment							
Quality of life	Ecological momentary assessment							
Physical activity	Mobile sensing		Continuously during the entire intervention period					
Smartphone usage	Mobile sensing							
Phone calls	Mobile sensing							
Text messages	Mobile sensing							

*The following subscales of the SCL-90 were used: organic syndrome, anxiety, insomnia, aggression, compulsion, pain, somatic symptoms and trauma.

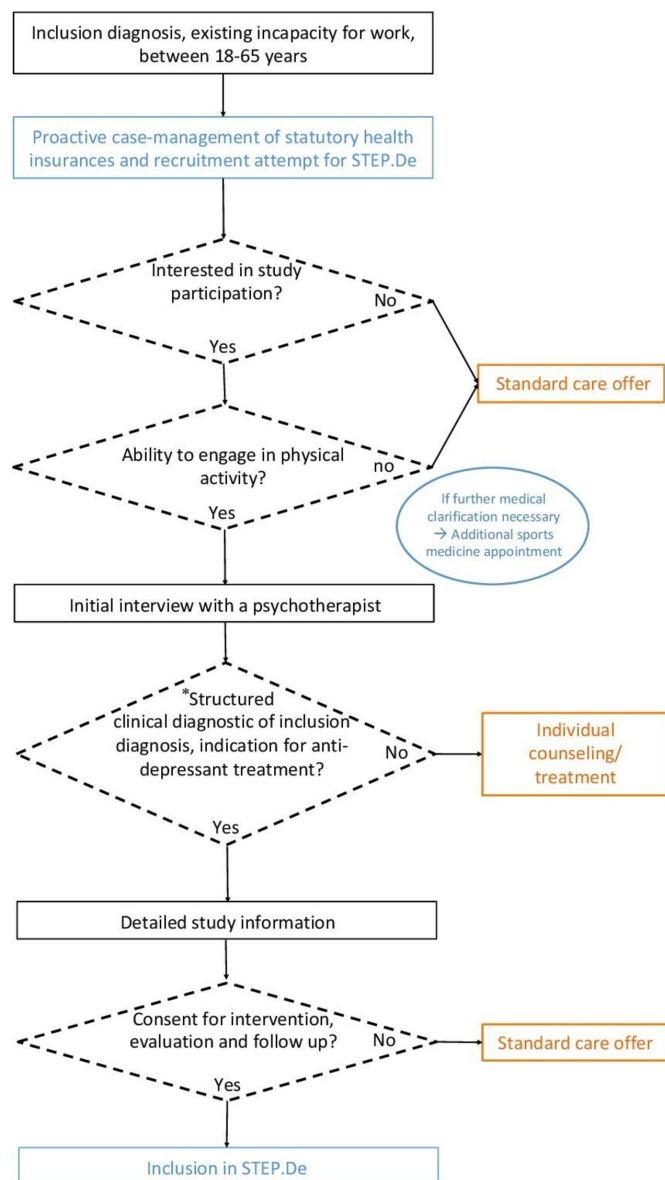


Figure 2 Recruitment flow chart. *Structured diagnoses are verified in a subgroup, because SCID documentation is not available for evaluation. SCID, Structured Clinical Interview for DSM; STEP.De, Sport/ Exercise Therapy for Depression.

tables 3, 4). Informed written consent will be obtained from all participants.

Participant eligibility

The study will recruit 480 outpatients (aged 18–65 years, both female and male). Patients will be recruited through four local health insurance carriers in Berlin, Germany. Case managers from the health insurance companies will contact insured people diagnosed with any of the inclusion diagnoses (F32.0, F32.1 (mild or moderate depressive episode), F33.0, F33.1 (recurrent depressive disorder, current episode mild or moderate), F34.1 (dysthymia), F43.2 (adjustment disorders), F43.8, F43.9 (reaction to severe stress), F48.0 (neurasthenia) and F41.2 (mixed anxiety and depressive disorder)) according to a general practitioner in combination with an existing incapacity

for work and ask whether they are interested in information regarding a programme to relieve depression. If they agree, the STEP.De study and programme will be introduced and some exclusion criteria will be requested. Based on a predefined list of consecutive random numbers (free permutation two codes (in the ratio 2:1) within blocks of 120 numbers), the case manager will then allocate the patient to one of the two treatment conditions (exercise or psychotherapy) and make an appointment with a psychotherapist from the respective IG. During an initial interview with the psychotherapist, indication for therapy is evaluated and eligibility criteria are confirmed: patients who are currently receiving ongoing outpatient psychotherapeutic treatment or long-term care or are under legal guardianship will be excluded from participating in the project. Additionally, patients with physical disabilities for whom exercise is not suitable and patients with other serious mental or neurological illnesses or current severe substance use disorders, as well as patients on long-term medication with benzodiazepines, opiates (several weeks), or high-dose (>70% of the maximum daily dose) pharmacotherapy (tricyclics, antipsychotics and neuroleptics) will also be excluded. A further inclusion criterion is the ability to engage in exercise therapy according to an adjusted German version of the Physical Activity Readiness Questionnaire.⁴⁴ In addition to the practitioner's diagnoses, psychotherapists provide a second clinical diagnosis based on their interview. Psychotherapists are instructed to use the Structural Clinical Interview I for Diagnostic and Statistical Manual of Mental Disorders 4 (DSM IV), Structured Clinical Interview for DSM (SCID) I, Axis 1, Section A, E and I. A subgroup of patients will be asked for a (second) SCID interview performed by trained undergraduates to evaluate consistency (see figure 2 for the detailed recruitment process).

Interventions

Exercise therapy participants will take part in a 16-week multicomponent exercise programme supervised by specifically trained exercise professionals (sport scientists, exercise therapists or physiotherapists with a bachelor's or higher degree in sport sciences or equivalent). Therapy group meetings take place twice a week, each session lasting 60 min. Patients can participate in a maximum of 32 exercise sessions over the treatment period. The two centres that are conducting the exercise therapy in the study have the same framework conditions (eg, same equipment) for providing exercise therapy. The number of participants per group ranges from 4 to 12 participants, in such a way as to enable the exercise professionals' individual instruction and supervision. Exercise professionals are required to be experienced in the specific field of exercise therapy/health exercise training (minimum of 3 months of professional experience). Additionally, they will receive a qualification training (16 units of training related to mental disorders with a focus on depression and emergency management) and a psychological qualification training (16 units of

training related to basic psychological needs support, motivational communication). The exercise therapy regimen is based on the following guidelines and empirical findings: the design of the aerobic exercise intervention follows the official guidelines of the American Heart Association⁴⁵ for aerobic exercise. Furthermore, recommendations from comparative meta-analyses in MDD^{13–15} regarding the type, duration, intensity, frequency and content of exercise are taken into consideration. The exercise therapy starts with an entrance test implemented by the exercise professionals. The entrance test includes a physical history (mainly orthopaedic) and physical tests (balance analysis, strength and core stability testing) to enable the implementation of an individual training schedule. Each exercise therapy session is divided into three parts: 20 min of aerobic exercise with a bicycle ergometer or jogging/nordic walking, 20 min of strength training, and a flexible segment of 20 min including coordination, relaxation, flexibility or perception-based and mindfulness-based exercises (eg, myofascial training, breathing techniques and elements of yoga) chosen by the exercise professionals according to the specific needs of the group and the individual participant. Our research conclusions may be confounded to some extent as a result because mindfulness-based exercises have been shown to have significant effect on depression (see Introduction). However, we expect that the main effect of the exercise therapy programme is caused by the physical exercise itself and will be able to test this effect on a process level using objective accelerometry data. In addition, our results may also be influenced by other factors such as the supervision from the exercise professionals and the social interactions/contact among the patients during the intervention. Psychotherapists attend the exercise intervention period with 30 min telephone conversations once every 4 weeks to evaluate whether the current treatment still meets patients' individual needs or whether treatment change is needed (eg, crisis intervention, stationary treatment). After 16 weeks of treatment, the exercise therapy will be completed with a closing meeting led by an exercise professional to evaluate the goals that were set and prospects of including the exercise into the patients' daily lives. In a closing meeting with the psychotherapist, further psychotherapeutic care will be assessed and reviewed.

Over 16 weeks of intervention, participants in the psychotherapy group will receive psychotherapeutic treatment as usual in standard care. This includes preparatory sessions and psychotherapeutic treatment sessions. Sessions take place once a week and will last for 50 min. Mainly Cognitive Behavioral Therapy (CBT) will be implemented, but also other forms of psychotherapy that are approved by the state healthcare system (eg, psychodynamic therapy).

Sample size calculation

Sample size was calculated for a non-inferiority trial with parameters as follows: the non-inferiority margin was set

to three points on the Beck Depression Inventory (BDI-II) sum score at the time of post treatment which represents $0.3 \times \text{SD}$ (with $\text{SD}=10$, according to baseline data in the current study). This is close to but slightly higher than the minimal important difference of 0.24 suggested by Cuijpers *et al*⁶ for use in depression trials. The slight liberalisation is supported by similar margins of about 0.3 to $0.4 \times \text{SD}$ in other trials on activity-related treatment of depression (eg, $0.35 \times \text{SD}$ in Rhodes *et al*⁴⁶; $0.39 \times \text{SD}$ in Richards *et al*⁴⁷). Moreover, it accounts for the possibility that effect sizes in exercise therapy might specifically depend on individual patient indication and hence the criterion for non-inferiority was chosen not too conservative. Using the R package 'SampleSize4ClinicalTrials'⁴⁸ with a 2:1 ratio of participants in exercise ad psychotherapy IGs, a true effect of zero (non-inferiority), alpha error rate of 5% and 80% power, as well as non-inferiority margin of 3 and SD 10 (see above) a total of 312 patients will be needed to test for non-inferiority between groups at the time of post treatment. Accounting for additional 25% of dropout, $n=390$ are planned to be included in the study. Note that intraclass correlation (ICC) is expected to be very small in psychotherapy⁴⁹ and particularly small in the present case, where patients in the exercise treatment condition are only seen once by their psychotherapists. However, a total of $n=480$ included patients (384 entering the analysis after dropout) would still account for a small ICC of 0.01 in clusters of max 24 patients each using the cluster correction and design effect as described by Donner and Klar.⁵⁰

Measures

An overview of the primary and secondary outcome measures and the assessment process is given in table 1. At each of the six measurement time points (t_0 – t_5), patients will undergo an online self-assessment using a specific online assessment programme for clinical trials (secu-Trial). At t_0 , t_2 , and t_4 , patients will additionally complete a telephone interview with external trained raters to assess work ability (WHO Disability Assessment Schedule (WHODAS) 2.0⁵¹), depressive symptoms (Hamilton Rating Scale for Depression (HAM-D)⁵²), psychopathological symptoms (Symptom Checklist 90 Scale (SCL-90)⁵³), physical activity (International Physical Activity Questionnaire (IPAQ)⁵⁴), previous depressive episodes (single item, assessed only at t_0), physical comorbidities (assessed only at t_0) and medication (assessed at t_0 , t_2 , t_4).

Primary outcome

The primary outcome is depression severity. The German version of the BDI-II⁵⁵ will be used to assess the severity of depression.

Secondary outcomes

Secondary outcomes include the depressive symptoms (assessed by the HAM-D,⁵⁶ t_0 , t_2 , t_4 , via external telephone rating), work ability (assessed by the Work and

Social Adjustment Scale⁵⁷, t0–t5, and WHODAS 2.0,⁵¹ t0, t2, t4, via telephone assessment), psychopathological symptoms (assessed by predefined subscales and items of the SCL-90,⁵³ t0, t2, t4 via telephone assessment), psychological needs satisfaction and frustration (assessed by the Basic Psychological Need Satisfaction and Frustration Scale,⁵⁸ t0, t2, t4), and health-related quality of life (assessed by the EuroQol Five-Dimensional Questionnaire,⁵⁹ t0–t5, and the SF-12,⁶⁰ t0, t2, t4). Physical activity (assessed by the short form of the IPAQ,⁶¹ t0, t2, t4), self-efficacy (assessed by the General Self-Efficacy Scale,⁶² t0, t2, t4), healthcare climate (assessed by the Health Care Climate Questionnaire,⁶³ t0, t1, t2), sports motivation (assessed by the 12-Item Short Form Behavioral Regulation Questionnaire,⁶⁴ t0, t2, t4) and mindfulness (assessed by the Mindfulness Attention Awareness Scale,⁶⁵ t0, t2, t4) will be assessed in the present study as process variables.

Cost efficiency

Health insurance cost data will be collected to assess the cost efficiency. Costs per case, that is, expenditures in the 12 months before and after the intervention will be included. Selected additional covariates will be used to carry out exploratory subgroup analyses on the sustainability of the two interventions. For the subgroup analyses, the progress of further psychotherapy (degree, short-term therapy, long-term therapy, group therapy) will be categorically assessed, as well as the specific diagnosis in each case.

Process evaluation

Evening assessments of the continuous EMA include three depressive symptoms (mood, anhedonia, fatigue/loss of energy), frequency and success of four emotion regulation strategies (reappraisal, suppression, distraction, rumination), three categories of physical activity (low, moderate, high) and three questions regarding cognitive control (organisation, planning, concentration), as well as single items on adverse effects of the treatment, quality of life, self-efficacy, therapy motivation and progress. Higher sampled EMA will focus on five emotional states (anger, sadness, joy, anxiety, guilt), emotion regulation strategies and cognitive control. All items will be presented in a pseudorandom blockwise order (eg, emotional regulation strategies will be queried in pseudorandom order but always together). The n-back paradigm will include 4 blocks of 2-back and 3-back each, as well as 2 blocks of 0-back (control condition). Blocks will be presented in pseudorandom order (no more than two repetitions) and include 5 targets and 11 non-target trials (stimulus duration 500 ms, intertrial interval 1200 ms). Mobile sensing will include continuous assessment of physical activity (using Google's activity recognition API) as well as phone usage, call and text message activity. Objective Assessment of physical activity will be based on accelerometer data measured with GENEActive devices at 50 Hz with a minimum wear time of 70%, and mainly focus on the Euclidian Norm Minus One (ENMO) metric. Due

to individual potential overtaxing, the process evaluation is not a requirement to participate in the study. Also, the outcome evaluation can be reduced to a minimum on request from the participants at any time.

Planned statistical analyses

All statistical analyses will be carried out on an intention-to-treat basis. Thereby, all participants that completed baseline data are included in the analyses with the last observation carried forward. If people drop out, we call them and ask for the main outcome at that moment. The analysis of the primary outcome will be blinded to ensure the independence of the evaluation (ie, the statistician will be blinded as to which group signifies IG and TAU). For the primary endpoint (severity of depression measured by BDI-II), linear regression analyses with cluster adjustment will be used. To account for the dependence of the measurements within a cluster, generalised linear estimation equations will be used. For secondary endpoints, mixed linear and logistic regression models with fixed and random factors will be added to estimate cluster heterogeneity. Values at baseline and differences in baseline data are considered as confounders in the model estimates. For health economic analyses, costs per case will be determined in a difference-in-difference approach from the perspective of the statutory health insurance. To determine the cost efficiency, expenditures in the 12 months before and after the intervention and the outcome difference (improvement on the BDI-II) will be compared between groups in a multilevel model. Continuous time models⁶⁶ and linear mixed models will be applied to analyse process data from EMA, mobile sensing and accelerometry over 16 weeks.

Key methodological issues

Key methodological issues are: (1) the use of a CRCT design, (2) implementation of a pragmatic RCT and (3) application of a non-inferiority trial.

Cluster RCT

Since the STEP.De study is aimed not only at patients but also at units within the health services that involve the restructuring of healthcare delivery, the implementation of a CRCT rather than an individual-level RCT is preferred.⁶⁷ The CRCT allows the study of relevant effects at the level of health services (at the level of the psychotherapists, in this case). It can be used to determine the comparative effectiveness of two or more therapy options under conditions of actual use by incorporating existing clinical and administrative variables (eg, existing information about the participants, treatments and outcomes, along with existing service provision data). This approach offers considerable cost and time efficiency by simplifying the logistics of implementation.⁶⁸ Importantly, CRCTs also circumvent the problem of the treatment effect being dependent on the skills of the individual health professional delivering it and reduces the risk of contamination.^{69 70} However, compared with the individual-level

RCT, a CRCT is more difficult to perform, which requires more participants to obtain the same statistical power and necessitates more complex statistical analysis⁴⁰ (eg, adjustment for the ICC coefficient of the cluster randomisation). Thus, in the present study, a relatively large sample of participants will be recruited to address these stipulations and a more rigorous data analysis strategy will be implemented (as discussed in the Planned statistical analyses section).

Pragmatic RCT

Taking into account the important limitation of efficacy trials that they are typically too far from usual conditions and therefore their external validity is questionable,⁹ the STEP.De study will implement a pragmatic RCT to ensure that the methodological features will be as closely aligned with routine practice as possible. However, pragmatic trials with high external validity may face the pitfall of compromising their internal validity. Therefore, a fine balance will be needed between interfering as little as possible with routine practice while still adhering to the methodological requirements of a rigorous RCT research design. Furthermore, contrary to traditional RCTs that typically run double blind, the STEP.De study will be an open-label study. In a pragmatic trial, therapist and patient biases are not necessarily viewed as detrimental but are rather accepted as part of the therapists' and patients' responses to the treatment and are included in the overall assessment.⁷¹ In the present study, all involved parties including patients, psychotherapists, exercise professionals and researchers expect both treatments to be equally helpful in reducing depressive symptoms. Therefore, the placebo effects will be comparably distributed between both conditions.

Non-inferiority trial

The implementation of a non-inferiority trial in the STEP.De study aims to demonstrate that the difference between exercise and psychotherapy treatment modalities is small enough (ie, lower than a predefined acceptable margin) to support the conclusion that exercise therapy is as effective a treatment strategy for patients with depressive symptoms as psychotherapy. However, the use of non-inferiority trials has received some critique over the years on the grounds that they merely study a new marketable treatment without offering any advantages over the existing treatment,⁷² and 'they disregard patients' interests, are therefore unethical,⁷³ because there needs to be an effective treatment that provides a benefit to the patients. In the current study, the non-inferiority of the exercise treatment compared with psychotherapy treatment is of interest on the premise that exercise treatment has some advantages over psychotherapy, such as greater availability, lower thresholds, quicker accessibility, relatively lower costs and benefits for both psychological and somatic health conditions. An active control (ie, psychotherapy as usual care) will be used in the present trial to address the ethical concern of using a placebo or a no-treatment

control group. It will be ensured that the psychotherapy control group receives TAU and, at the same time, the designed exercise therapy regimen will be based on treatments that have been proven to be effective to ensure that all participants receive adequate treatment to prevent any ineffective, unsafe or inferior treatments. Compared with the majority of previous studies, the high quality of the individually tailored, structured and supervised exercise therapy from the qualified exercise professionals features one of the main characteristics of the present study (for a detailed description, see the Interventions section). Nevertheless, after the exercise therapy, the psychotherapists allocated to the exercise group will decide whether further psychotherapeutic treatment (based on clinical assessment of treatment need) is needed for the exercise group participants to ensure treatment without a gap.

Patient and public involvement

This study protocol was written without patient involvement. Patients were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or interpret the future results. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy. However, we did include acceptability and tolerability data from an ongoing RCT of a comparable exercise intervention in patients with depressive disorder.⁷⁴

ETHICS AND DISSEMINATION

The trial will be conducted in compliance with this study protocol and the Declaration of Helsinki. Participation in the study will be on a voluntary basis. Informed written consent will be obtained from all participants and study-related materials will be handled correctly. All the professionals involved in the study will undergo specific trainings. The study will be reported in accordance with the CONSORT statement, especially its adapted extensions for pragmatic trials, CRCTs and non-inferiority trials. The results will be published in peer-reviewed academic journals and disseminated to the public.

Contributors AH, SH and MR designed the trial. AH, MR, SH and AP developed the intervention. MS, GW and SH contributed significantly to the trial formation. AH prepared the first draft of the manuscript and KA contributed to the first draft of the manuscript, followed by extensive revision by MR, SH, GW and AH. All authors contributed to and approved the final manuscript.

Funding This work was funded by the Innovationsausschuss des Gemeinsamen Bundesausschusses (G-BA; NVF1_2017-136; Innovation Fund of the Joint Federal Committee), and by an intramural junior research group grant to SH, MS and AH from the University of Potsdam. The study is partly funded by a grant from the University of Potsdam Graduate School to AH.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The study protocol was approved by the local ethics committee of the University of Potsdam, Potsdam, Germany (No. 17/2018) and the Freie Universität Berlin (No. 206/2018).

Provenance and peer review Not commissioned; externally peer reviewed.

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 iD

Andreas Heissel <http://orcid.org/0000-0001-9270-7027>

REFERENCES

- World Health Organization. *Depression and other common mental disorders: Global health estimates*. Geneva: World Health Organization, 2017.
- GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the global burden of disease study 2015. *Lancet* 2016;388:1545–602.
- Birnbaum HG, Kessler RC, Kelley D, et al. Employer burden of mild, moderate, and severe major depressive disorder: mental health services utilization and costs, and work performance. *Depress Anxiety* 2010;27:78–89.
- Belmaker RH, Agam G, Disorder MD. Major depressive disorder. *N Engl J Med Overseas Ed* 2008;358:55–68.
- Munder T, Flückiger C, Leichenring F, et al. Is psychotherapy effective? A Re-analysis of treatments for depression. *Epidemiol Psychiatr Sci* 2018;1–7.
- Cuijpers P, Karyotaki E, Reijnders M, et al. Was Eysenck right after all? A reassessment of the effects of psychotherapy for adult depression. *Epidemiol Psychiatr Sci* 2019;28:21–30.
- Kammer BP. Wartezeiten für psychisch kranke Menschen weiterhin zu Lang: BPTK zum G-BA-Gutachten Zur Bedarfsplanung, 2018. Available: <https://www.bptk.de/wartezeiten-fuer-psychisch-krankemenschen-weiterhin-zu-lang/> [Accessed 20 Mar 2019].
- Wittchen HU, Jacobi F, Rehm J, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol* 2011;21:655–79.
- Schuch FB, Morres ID, Ekkekakis P, et al. A critical review of exercise as a treatment for clinically depressed adults: time to get pragmatic. *Acta Neuropsychiatr* 2017;29:65–71.
- Lawlor DA, Hopker SW. The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. *BMJ* 2001;322:763.
- Schneider F, Härter M. *Schorr S. S3-Leitlinie/Nationale VersorgungsLeitlinie Unipolare Depression. Second*. Berlin: Springer, 2017. <https://www.leitlinien.de/nvl/depression/>
- Stubbs B, Vancampfort D, Hallgren M, et al. Epa guidance on physical activity as a treatment for severe mental illness: a meta-review of the evidence and position statement from the European psychiatric association (EPA), supported by the International organization of physical therapists in mental health (IOPTMH). *Eur Psychiatry* 2018;54:124–44.
- Morres ID, Hatzigeorgiadis A, Stathi A, et al. Aerobic exercise for adult patients with major depressive disorder in mental health services: a systematic review and meta-analysis. *Depress Anxiety* 2019;36:39–53.
- Schuch FB, Vancampfort D, Richards J, et al. Exercise as a treatment for depression: a meta-analysis adjusting for publication bias. *J Psychiatr Res* 2016;77:42–51.
- Cooney GM, Dwan K, Greig CA, et al. Exercise for depression. *Cochrane Database Syst Rev* 2013;CD004366.
- Honey EP. I shrunk the pooled SMD! guide to critical appraisal of systematic reviews and meta-analyses using the Cochrane review on exercise for depression as example. *Ment Health Phys Act* 2015;8:21–36.
- Heinzel S, Lawrence JB, Kallies G, et al. Using exercise to fight depression in older adults: a systematic review and meta-analysis. *GeroPsych* 2015;28:149–62.
- Nebiker L, Lichtenstein E, Minghetti A, et al. Moderating effects of exercise duration and intensity in neuromuscular vs. endurance exercise interventions for the treatment of depression: a Meta-Analytical review. *Front Psychiatry* 2018;9:305. Epub ahead of print.
- Rethorst CD, Wipfli BM, Landers DM. The antidepressive effects of exercise. *Sports Medicine* 2009;39:491–511.
- Gordon BR, McDowell CP, Hallgren M, et al. Association of efficacy of resistance exercise training with depressive symptoms: meta-analysis and meta-regression analysis of randomized clinical trials. *JAMA Psychiatry* 2018;75:566–76.
- Cramer H, Lauche R, Langhorst J, et al. Yoga for depression: a systematic review and meta-analysis. *Depress Anxiety* 2013;30:1068–83.
- Giles GE, Cantelon JA, Eddy MD, et al. Habitual exercise is associated with cognitive control and cognitive reappraisal success. *Exp Brain Res* 2017;235:3785–97.
- Chang Y-K, Huang C-J, Chen K-F, et al. Physical activity and working memory in healthy older adults: an Erp study. *Psychophysiology* 2013;50:1174–82.
- Schweizer S, Grahn J, Hampshire A, et al. Training the emotional brain: improving affective control through emotional working memory training. *J Neurosci* 2013;33:5301–11.
- Xiu L, Zhou R, Jiang Y. Working memory training improves emotion regulation ability: evidence from HRV. *Physiol Behav* 2016;155:25–9.
- Stanton R, Reaburn P. Exercise and the treatment of depression: a review of the exercise program variables. *J Sci Med Sport* 2014;17:177–82.
- Hallgren M, Helgadóttir B, Herring MP, et al. Exercise and Internet-based cognitive-behavioural therapy for depression: multicentre randomised controlled trial with 12-month follow-up. *Br J Psychiatry* 2016;209:414–20.
- Mota-Pereira J, Silvério JMA, Fonte D, et al. Moderate exercise as an adjuvant therapy for treatment-resistant major depressive disorder: 6-month follow-up. In: Simmons JA, Brown AC, eds. *Aerobic exercise: health benefits, types and common misconceptions*. New York: Nova Science Publishers, 2013. <https://repositorium.sdum.uminho.pt/>
- Helgadóttir B, Forsell Y, Hallgren M, et al. Long-Term effects of exercise at different intensity levels on depression: a randomized controlled trial. *Prev Med* 2017;105:37–46.
- de Maat SM, Dekker J, Schoevers RA, et al. Relative efficacy of psychotherapy and combined therapy in the treatment of depression: a meta-analysis. *Eur Psychiatry* 2007;22:1–8.
- Shinohara K, Honyashiki M, Imai H, et al. Behavioural therapies versus other psychological therapies for depression. *Cochrane Database Syst Rev*;138.
- Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6:e1000100.
- Stubbs B, Vancampfort D, Rosenbaum S, et al. Challenges establishing the efficacy of exercise as an antidepressant treatment: a systematic review and meta-analysis of control group responses in exercise randomised controlled trials. *Sports Med* 2016;46:699–713.
- Bower P, Byford S, Sibbald B, et al. Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression. II: cost effectiveness. *BMJ* 2000;321:1389–92.
- Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new medical Research Council guidance. *BMJ* 2008;337:a1655.
- Burcusa SL, Iacono WG. Risk for recurrence in depression. *Clin Psychol Rev* 2007;27:959–85.
- Chan A-W, Tetzlaff JM, Gøtzsche PC, et al. Spirit 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;346:e7586.
- Schulz KF, Altman DG, Moher D, et al. Consort 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Med* 2010;8:18.
- Zwarenstein M, Treweek S, Gagnier JJ, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;337:a2390.
- Campbell MK, Piaggio G, Elbourne DR, et al. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012;345:e5661.
- Piaggio G, Elbourne DR, Pocock SJ, et al. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2012;308:2594.
- Loudon K, Treweek S, Sullivan F, et al. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;350:h2147.
- Maher CG, Sherrington C, Herbert RD, et al. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther* 2003;83:713–21.
- Finger JD, Gößwald A, Härtel S, et al. Messung Der kardiorespiratorischen fitness in Der studie Zur Gesundheit Erwachsener in Deutschland (DEGS1). *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2013;56:885–93.
- American Heart Association. American heart association recommendations for physical activity in adults and kids, 2018. Available: www.heart.org, <https://www.heart.org/en/healthy-living/>



- fitness/fitness-basics/aha-recs-for-physical-activity-in-adults [Accessed 4 Mar 2019].
- 46 Rhodes S, Richards DA, Ekers D, *et al.* Cost and outcome of behavioural activation versus cognitive behaviour therapy for depression (cobra): study protocol for a randomised controlled trial. *Trials* 2014;15:29.
- 47 Richards DA, Ekers D, McMillan D, *et al.* Cost and outcome of behavioural activation versus cognitive behavioural therapy for depression (cobra): a randomised, controlled, non-inferiority trial. *Lancet* 2016;388:871–80.
- 48 Qi H, Zhu F. Sample size calculation for the comparison of means or proportions in phase III clinical trials, 2019. Available: <https://cran.r-project.org/web/packages/SampleSize4ClinicalTrials/SampleSize4ClinicalTrials.pdf> [Accessed 5 Dec 2019].
- 49 Kuyken W, Byford S, Taylor RS, *et al.* Mindfulness-based cognitive therapy to prevent relapse in recurrent depression. *J Consult Clin Psychol* 2008;76:966–78.
- 50 Donner A, Klar N. Cluster randomization trials in epidemiology: theory and application. *J Stat Plan Inference* 1994;42:37–56.
- 51 Ustün TB, Chatterji S, Kostanjsek N, *et al.* Developing the world Health organization disability assessment schedule 2.0. *Bull World Health Organ* 2010;88:815–23.
- 52 Schmitt A, Kulzer B, Hermanns N. Deutsche Fassung grid Hamilton rating scale für depression (GRIDHAMD) 2012. Epub ahead of print.
- 53 Franke G. *Die Symptom-Check-Liste von Derogatis: Deutsche Version*. Göttingen: Beltz Test Gesellschaft, 1995.
- 54 Craig CL, Marshall AL, Sjöström M, *et al.* International physical activity questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise* 2003;35:1381–95.
- 55 Hautzinger M, Keller F, Kühner C. *Das Beck Depressions Inventar II: Deutsche Bearbeitung und Handbuch zum BDI II*. Frankfurt: Harcourt Test Services, 2006. <https://www.testzentrale.de/shop/beck-depressions-inventar-revision.html>
- 56 Hamilton M. A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry* 1960;23:56–62.
- 57 Mundt JC, Marks IM, Shear MK, *et al.* The work and social adjustment scale: a simple measure of impairment in functioning. *Br J Psychiatry* 2002;180:461–4.
- 58 Heissel A, Pietrek A, Flunger B, *et al.* The validation of the German basic psychological need satisfaction and frustration scale in the context of mental health. *European Journal of Health Psychology* 2018;25:119–32.
- 59 Herdman M, Gudex C, Lloyd A, *et al.* Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727–36.
- 60 Gandek B, Ware JE, Aaronson NK, *et al.* Cross-Validation of item selection and scoring for the SF-12 health survey in nine countries. *J Clin Epidemiol* 1998;51:1171–8.
- 61 Lee PH, Macfarlane DJ, Lam TH, *et al.* Validity of the International physical activity questionnaire short form (IPAQ-SF): a systematic review. *Int J Behav Nutr Phys Act* 2011;8:115.
- 62 Romppel M, Herrmann-Lingen C, Wachter R, *et al.* A short form of the general self-efficacy scale (GSE-6): development, psychometric properties and validity in an intercultural non-clinical sample and a sample of patients at risk for heart failure. *GMS Psycho-Soc-Med* 2013;10:1–7.
- 63 Heissel A, Pietrek A, Rapp MA, *et al.* Perceived health care climate of older people attending an exercise program: validation of the German short version of the health care climate questionnaire. *J Aging Phys Act* 2019:1–11–11.
- 64 Kleinert J, Wasserkampf A. *A short and context-free behavioral regulation questionnaire (BRQ-12)*. Köln: Deutsche Sporthochschule Köln, 2016.
- 65 Brown KW, Ryan RM. The benefits of being present: mindfulness and its role in psychological well-being. *J Pers Soc Psychol* 2003;84:822–48.
- 66 Voelkle MC, Oud JHL. Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes. *Br J Math Stat Psychol* 2013;66:103–26.
- 67 Gilbody S, Whitty P. Improving the delivery and organisation of mental health services: beyond the conventional randomised controlled trial. *Br J Psychiatry* 2002;180:13–18.
- 68 Velentzas P, Mohr P, Messner DA. *Making informed decisions: Assessing the strengths and weaknesses of study designs and analytic methods for comparative effectiveness research*. Washington, DC: National Pharmaceutical Council, 2012.
- 69 Walters SJ. Therapist effects in randomised controlled trials: what to do about them. *J Clin Nurs* 2010;19:1102–12.
- 70 Hotopf M. The pragmatic randomised controlled trial. *Adv. psychiatr. treat* 2002;8:326–33.
- 71 Roland M, Torgerson DJ. Understanding controlled trials: what are pragmatic trials? *BMJ* 1998;316:285.
- 72 Pocock SJ. The pros and cons of noninferiority trials. *Fundam Clin Pharmacol* 2003;17:483–90.
- 73 Garattini S, Bertele V. Non-Inferiority trials are unethical because they disregard patients' interests. *Lancet* 2007;370:1875–7.
- 74 Heinzel S, Rapp MA, Fydrich T, *et al.* Neurobiological mechanisms of exercise and psychotherapy in depression: the speed study-Rationale, design, and methodological issues. *Clin Trials* 2018;15:53–64.