Available online at www.sciencedirect.com
ScienceDirect
Contemporary Clinical Trials 28 (2007) 79–89
Contemporary Clinical Trials
www.elsevier.com/locate/conclintrial

Design paper: The DEMO trial: A randomized, parallel-group, observer-blinded clinical trial of aerobic versus non-aerobic versus relaxation training for patients with light to moderate depression

Jesper Krogh [a,b,c,*], Lone Petersen [a], Michael Timmermann [b], Bengt Saltin [b], Merete Nordentoft [a]

[a] Psychiatric Department E, Bispebjerg Hospital, Copenhagen University Hospital,Copenhagen, Denmark
[b] Copenhagen Muscle Research Center, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
[c] Copenhagen Trial Unit, Center for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

Received 15 August 2005; accepted 5 July 2006

Abstract

Background: In western countries, the yearly incidence of depression is estimated to be 3–5% and the lifetime prevalence is 17%. In patient populations with chronic diseases the point prevalence may be 20%. Depression is associated with increased risk for various conditions such as osteoporoses, cardiovascular diseases, and dementia. WHO stated in 2000 that depression was the fourth leading cause of disease burden in terms of disability. In 2000 the cost of depression in the US was estimated to 83 billion dollars. A predominance of trials suggests that physical exercise has a positive effect on depressive symptoms. However, a meta-analysis from 2001 stated: "The effectiveness of exercise in reducing symptoms of depression cannot be determined because of a lack of good quality research on clinical populations with adequate follow-up."
Objectives: The major objective for this randomized trial is to compare the effect of non-aerobic, aerobic, and relaxation training on depressive symptoms using the blindly assessed Hamilton depression scale ($HAM\text{-}D_{17}$) as primary outcome. The secondary outcome is the effect of the intervention on working status (i.e., lost days from work, employed/unemployed) and the tertiary outcomes consist of biological responses.
Design: The trial is designed as a randomized, parallel-group, observer-blinded clinical trial. Patients are recruited through general practitioners and psychiatrist and randomized to three different interventions: 1) non-aerobic, — progressive resistance training, 2) aerobic training, — cardio respiratory fitness, and 3) relaxation training with minimal impact on strength or cardio respiratory fitness. Training for all three groups takes place twice a week for 4 months. Evaluation of patients' symptoms takes place four and 12 months after inclusion. The trial is designed to include 45 patients in each group.

Statistical analysis will be done as intention to treat (all randomized patients).

Results from the DEMO trial will be reported according to the CONSORT guidelines in 2008–2009.

Keywords: Depression; Exercise; Aerobic; Non-aerobic; Work; Training

* Corresponding author. Psychiatric Department E, Bispebjerg Hospital, Copenhagen University Hospital, DK-2400, Copenhagen, Denmark. Tel.: +45 3531 6682; fax: +45 3531 3953.
E-mail address: jk27@bbh.hosp.dk (J. Krogh).

1551-7144/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.cct.2006.07.001

1. Introduction

The yearly incidence of depression is estimated to be 3–5% [1–3], with a lifetime prevalence of 17% in western societies [2]. The prevalence of major depression in patients with chronic and disabling physical illnesses can be as high as 20% [4]. Depressive symptoms as prolonged feeling of sadness, low self esteem, and even suicidal tendencies have consequences to the patients' social life, and the severity of depression correlate to the number of work days lost [5,6]. Depression is associated with increased risk for conditions such as osteoporoses [7,8], cardiovascular diseases [9–12], and dementia [13]. WHO and others stress that unipolar depressive disorders are the fourth leading cause of global disease burden in terms of lost years of healthy life and that major depression accounted for 12% of all years lived with disability in 2000 [14]. The economic burden of depression on the national economy in the US was in 2000 estimated to 83.1 billion dollars (62% were workplace costs, 31% were direct medical costs, and 7% were suicide-related mortality costs) [15].

A meta-analysis of randomized trials comparing either non-aerobic or aerobic training with no intervention in patients diagnosed with depression concluded that the effect of training cannot be determined due to lack of good quality research [16]. The authors found that the majority of trials were without blinded outcome assessment, lacked intent-to-treat analyses, and most had short follow-up. This conclusion is supported in a later review [17]. The methodological problems mentioned above might lead to an overestimation of intervention effect [18–20]. Two recently published trials, comparing aerobic training versus no treatment [21] and non-aerobic training versus standard general practitioner treatment [22], suggest that training has a positive effect on patients diagnosed with depression. Furthermore they found the positive effect of training to be intensity related. Only one previous trial has compared non-aerobic and aerobic training in the same trial [23]. The non-aerobic and aerobic training could theoretically be influencing depressive symptoms by different biological mechanisms, such as enhanced serotonergic activity due to enhanced free tryptophan levels in aerobic exercise [24,25].

On this background we designed a randomized trial comparing non-aerobic training versus aerobic training versus a control group receiving relaxation training. This trial will employ adequate methods of reducing bias such as centralized randomization, blinded outcome assessment, and analysis according to the intent-to-treat principle. Furthermore we will include patients diagnosed with light to moderate depression and have a long follow-up to evaluate the benefits and harms of training on depressive symptoms, lost days from work, and biological outcome measures.

2. Objectives for the DEMO trial

The primary objective for the DEMO superiority trial is to assess the beneficial and harmful effects of non-aerobic versus aerobic versus relaxation training in patients diagnosed with depression according to ICD-10. The primary outcome will be an assessment of depressive symptoms as measured with the Hamilton Depression scale (HAM-D_{17}) [26] by outcome assessors blinded for training group.

3. Design

The trial is designed as a three-armed, parallel-group, observer-blinded randomized clinical superiority trial (Fig. 1). The patients are randomized to 1) a non-aerobic exercise program, 2) an aerobic exercise program, and a 3) relaxation exercise program. The relaxation group functions as a control group with minimal muscle contractions, and is established in order to eliminate the potential positive effect of being in a group on a regular basis.

3.1. Recruitment and inclusion and exclusion criteria

Physicians and the public are informed of the purpose of the DEMO trial and the inclusion and exclusion criteria through open meetings, newspapers, leaflets, and television. Patients are referred from general practitioners, private practicing psychiatrists, psychologist, and psychiatric wards to the first interview, which determines if the patient fulfills the inclusion criteria and none of the exclusion criteria (Table 1).

The age limits (18–55 years) is chosen in order to recruit patients likely to be active on the labor market, and to limit the amount of physical testing needed before exercise intervention can be started. In order to increase adherence to the randomized intervention we want the patients to live in the vicinity of the exercise facilities. Patients with suicidal behavior at entry (defined as a score higher than two on HAM-D_{17} item 3) are excluded from the trial and referred to

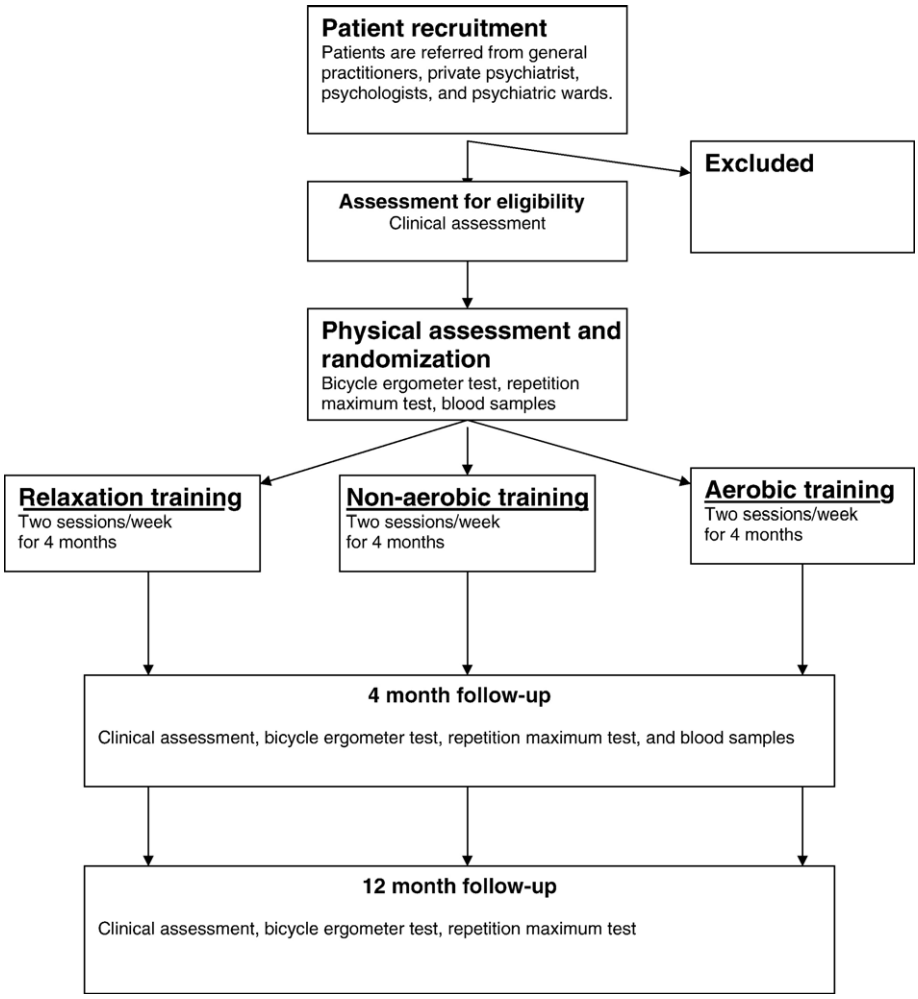

Fig. 1. Flowchart of the DEMO trial. See text for further information.

other treatment. Patients with sick leave for longer than 24 months are excluded since they are considered difficult to get back to work, and therefore the effect on sick leave for shorter duration might be missed. Following written and verbal informed consent and the clinical assessment and physical test described below, the patients are randomized.

3.2. Clinical assessment

A research assistant and an experienced psychologist conduct the interview of the participants. The diagnosis of depression is established using Major Depression Inventory (ICD-10) [27] and the severity of depression by HAM-D_{17}

Table 1
Inclusion and exclusion criteria for the DEMO trial

Inclusion:	Exclusion:
1. Age: 18 to 55 years	1. Current substance or recreational drug abuse
2. Living in the vicinity of Copenhagen	2. Contraindications to physical exercise
3. Depression (F32.0, F32.1, F33.0, ICD-10 verified F33.1)	3. Exercising more than 1 h/week
4. Fluent in Danish	4. Sick leave >24 months
5. Able to read and understand informed consent	5. >2 on item 3 (HAM-D_{17})
	6. Early-retirement

[26]. Several other psychometrics are used in order to monitor differences in different scales (i.e., interviewer versus self-administered scales) and to secure sufficient data on psychopathology and cognitive function (Fig. 2). Both interviewers are trained and certified in using the HAM-D_{17} interview. Inter-rater reliability will be calculated. The patient group assignment is unknown to the interviewers. In relation to the follow-up interview, the patients are instructed not to reveal their group assignment. At the follow-up the interviewer will be asked to 'guess' which group the patient had been assigned to, making it possible to reveal if the 'blinding' is unsuccessful.

3.3. Physical exercise testing

The patient's maximal oxygen uptake ($V{\cdot}O_{2max}$) is measured using a bicycle cardiopulmonary exercise test. The test is based on L. B. Andersens cycle exercise protocol [28]. The patients are instructed to fast from midnight, and they meet between 8:30 and 10:0 a.m. In the initial 5 min of the cycle test (Monark) the workload is 75 W for women, and 100 W for men. Then the workload is increased by 25 W/2 min till exhaustion. All patients are continuously verbally encouraged. The pulse and perceived level of exhaustion (PLE) on the Borg 15 graded category scale of the patients are recorded after 5 min, +50 W, +100 W, and at $V{\cdot}O_{2max}$.

An indwelling venous catheter is inserted in the antecubital vein and blood samples are collected 15 min prior to exercise testing (after 5 minutes rest), at $V{\cdot}O_{2max}$, and at 15, 30, and 60 min after the exercise test. All samples are collected in a sitting position. Blood samples that are not immediately analyzed are centrifuged and the plasma stored at − 80 °C until analyzed.

Repetition maximum (RM) for knee extension, chest press, and leg press is measured using machines and standardized procedures.

3.4. Randomization

If the inclusion criteria are fulfilled and none of the exclusion criteria is fulfilled the patients are subsequently randomized. Randomization is done as restricted randomization with the patients being stratified for: 1) not receiving antidepressive medication, 2) having received antidepressive medication for less than 6 weeks, and 3) having received antidepressive medication for more than 6 weeks. The rationale for this is the significant effect of antidepressive

Psychopathology:

Hamiltons Depression Scale[26] (primary outcome)
Montgomery-Asberg Depression Rating Scale[42]
Becks Depression Inventory II[43]
Hamiltons Anxiety Scale[44]
Symptom Checklist[76]

Cognitive test:

Trailmaking A and B[53]
Danish Adult Reading Test [50]
Verbal Fluency[54]
Substracting Serial Sevens [52]
WAIS- Digit Span[77]
Buschke Selective Reminding Test[55]
Rey-Osterrieth Complex Figure Test [56]
WAIS- R Digit Symbol Test[51]

Biochemistry

Cortisol
Prolactin

Quality of life:

WHO-5[27]

Fig. 2. Instruments and para-clinics used in the DEMO trial [76,77].

medication within 6 weeks [29]. The randomization is centralized and carried out at the Copenhagen Trial Unit. The block size is unknown to the investigators of the DEMO trial.

3.5. Interventions

All three intervention groups meet twice a week in the afternoon for 4 months at the same location. The patients can choose between a group that starts at 2 p.m. or 5 p.m. The patients are responsible for transportation themselves. All groups have 20 min of warm-up but with low intensity in the relaxation group. The whole training session is scheduled to last one and a half hour. The patients in the aerobic and relaxation training groups carry pulse monitor (Polar m-31 and m-61). In the aerobic exercise group it is used continuously to secure that the patients are exercising in the prescribed pulse interval. In the relaxation group it is used a few times during the 4 months to secure a low pulse (maximum 120; Borg 12).

3.5.1. Non-aerobic training group

This program is designed to increase muscular strength as measured by repetition maximum (RM) with initially 12 repetitions of 50% of RM 2–3 times per exercise (Table 2). As the participants progress the number of repetitions are reduced to 10 and 8 with an increase to 75% of RM. The training is a circle-training program involving large muscle groups including both machines and free weights. Six exercises are made on machines involving large muscle groups: Leg extension, leg press, total abdominal, lower back, chest press, vertical traction. As a supplement to this free weights and sandbags are used for exercising the calf muscle, the arm-abductors, the triceps brachi, and the hip abductors. The program fulfills the criteria for resistance training made by the American College of Sports Medicine (ACSM) [30] regarding number of repetitions and sets for untrained individuals. The loading is initially quite low (50% of RM), however, this loading has been noted to give an increase in strength in previously untrained individuals [31]. Two days training a week is described as a minimum, but still gives 80–90% of a three days/week program [32,33].

3.5.2. Aerobic training group

This program is designed to increase fitness as measured by maximal oxygen uptake ($V{\cdot}O_{2max}$) in a maximal cycle ergometer test. The exercise program involves ten different aerobic exercises using large muscle groups. Machines are used for cycling, running, stepping, abdominal exercises, and rowing. Additional exercises are sliding movements on small carpets, trampoline, step bench, jump rope, and Ski Fitter®. Initially each exercise lasts 2 min, followed by two minutes rest, and repeated twice which amounts to a total exercise time of 40 min. This gradually increases to three minutes exercise and one-minute rest, and each exercise repeated twice amounting to a total exercise time of 60 min. When patients become acclimated to the training the intensity rises (Table 3).

We chose training twice a week to increase adherence due to less injuries [34] and less motivation demanded from the patient, even though it could be argued that three days training/week as recommended by the ACSM [35] would have a higher impact on $V{\cdot}O_{2max}$. A recent study attributed the beneficial effect of exercise on depression to the intensity of training and found no differences comparing exercises 3 days/week and 5 days/week [21]. Earlier research indicates that duration and intensity are the most important factors in cardio respiratory fitness and that intensity levels of 60–80% of $V{\cdot}O_{2max}$ twice a week is sufficient to increase $V{\cdot}O_{2max}$ and comparable to training 3/week on this variable [36–38]. However training twice a week is probably not enough to change body composition [36,38]. The intensity and duration of the program fulfills the requirements for developing aerobic fitness according to ACSM [35] and the Danish health authorities [39]. Regarding the duration of each session the time ranged from 40 to 60 min of effective exercise time which should be sufficient to improve $V{\cdot}O_{2max}$.[34,40,41].

Table 2
Progression in the non-aerobic training group in the DEMO trial

Session numb.	RM[1]	Repetitions	Sets
1–4	20=50%	12	2–3
5–15	15=60%	12	3
16–20	12=70%	10	3
21–32	10=75%	8	4

[1]: Repetition maximum. The maximum weight that can be liftedonly one time.

Table 3
Intensity level for the aerobic training group in the DEMO trial

Session numb.	% of HR^{1}_{max}	Borg scale
1–8	70	12–13
9–23	80	14–15
24–32	89	15–16

[1]Maximum heart rate. Estimated during a cycle ergometer test.

3.5.3. Relaxation training group

In this group the goal is to avoid muscular contractions or stimulation of the cardiovascular system thus patients do not engage in activity perceived higher than 12 on the Borg scale, and the main exercises are done at an intensity level at 6–10 on the Borg scale. The first 20–30 min are used for movements on mattresses, Bobath balls®, or massage on the back using a Ball stick ball®. The massage is given for 10 min by another patient. This is followed by balance exercises and light exercises with tubes in a 10–20 minutes program. The sessions end with relaxation exercises for 20–30 min. The relaxation exercises are alternating muscle contraction and relaxation in different muscle groups while lying down.

4. Outcomes

All psychometric and biological outcomes are listed in Fig. 2, and explained below.

4.1. Psychometrics of depression

The HAM-D_{17} [26] is chosen as the primary outcome due to its widespread international use in psychiatric hospitals and among general practitioners. The scale consists of 17 items, some items can be scored from 0 to 2 and others from 0 to 4. The assessor will be blinded to intervention group. At the follow-up interviews, the patients are instructed not to reveal their group assignment. At the follow-up the interviewer will 'guess' which group the patient had been assigned to, making it possible to reveal if the 'blinding' is unsuccessful. In addition we include a depression rating scale which was designed to be more sensitive: The Montgomery Aasberg Depression Rating Scale [42] which is a 10-item scale. Each item can be rated from 0 to 6. The patient administered Becks Depression Inventory scale [43], which consists of 21 items that each can be scored from 0 to 3, is also included. Furthermore we include the 14-item Hamilton Anxiety Scale [44] to be able to evaluate the possible anxiety reducing effect of exercise.

4.2. Other psychometrics

To get a view of the patients psychiatric status we include the patient administered Symptom Questionnaire 92 (SCL-92) [45] which contains 92 items each rated from 0 to 4. On the basis of the answers the level of somatization, interpersonal sensitivity, depression, anxiety, phobic anxiety, obsession/compulsion, hostility, paranoid ideation, and psychoticism can be estimated. To get an estimate of the patients quality of life we include the patient administered 5-item WHO-5 scale where each item can be rated from 0 to 5 [27].

4.3. Lost days from work

At baseline and at four- and 12 months follow-up the patients are asked to report lost days from work during the last ten working days. Throughout the intervention the patients are asked to register these data as well.

4.4. Cognitive outcomes

New evidence for the biological effect of exercise includes the increase of brain derived neurotrophic factor, which is thought to mediate the positive effect of exercise on cognition in response to physical activity in rodents [46,47]. The deficits in cognitive function in depressed patients are widely recognized [48], and cognition has previously been shown

to relate to fitness levels in cross sectional studies [49]. The effect of long term physical intervention on cognitive skills has to our knowledge never been examined in clinically depressed patients.

4.4.1. Verbal intelligence

Danish Adult Reading Test [50] is a Danish version of the New Adult Reading Test, which consists of 50 irregular words graded in difficulty. The patient is asked to read out the list, and the number of correctly pronounced words is the score which has been interpreted as a measure of premorbid intelligence and correlates highly with a full-scale IQ measured by the Wechsler Adult Intelligence Scale [51].

4.4.2. Attention

Digit Span [51] makes the patient verbally repeat orally presented strings of digits of increasing lengths in straight and reversed order. The number of correctly repeated strings is the score. Subtracting Serial Sevens [52] requests the patients to subtract seven from 100 and continue to subtract seven until around zero. The score is a combination of time and number of errors.

4.4.3. Visuomotor speed

Visuomoter speed is tested with Trail Making part A and B [53] and the Digit Symbol Test [51]. In the A part, the patients are asked to connect numbered circles on a sheet in consecutive order. In the B part, the patients will have to connect numbers and letters in alternating sequence (A–1–B–2–C–3…). The patients are told to work as fast as possible. The score on each test is the time to complete. The Digit Symbol Test [51] is a symbol/number substitution test where the patients on top of a sheet of paper is presented with the numbers 1–9. Each number is represented by a symbol. Below there is 100 symbols listed but without the corresponding number. The patient is now asked to fill the corresponding numbers in 90 s. The number of correctly matched numbers and symbols is the score.

4.4.4. Language

Language is tested with Verbal Fluency S and Animals [54]. In the S part, the patients are asked to name as many words beginning with S as possible. They are not allowed to use proper nouns. In the Animal part, the patients are asked to mention as many animals they can think of. In both tests the patients have 60 s to name as many as possible. The number of correct and incorrect words is recorded for both tests.

4.4.5. Memory

In the Buschke [55] test a list of ten unrelated and different words are read aloud to the patient. The patient is then asked to recall the list. The interviewer repeats the words the patient misses and the patient is asked to try again until all ten words can be said, or until ten attempts.

After 10 min, where other tests is done, the patient is asked to recall as many of the words as possible. In the Rey Complex Figure Test [56] the patient is asked to copy a geometrical complex figure presented to them on a sheet of paper to another sheet of paper. When this is done the drawings and the original are put away and after 3 min the patient is asked to draw as much of the figure they can recall.

4.5. Biological outcomes

The potential beneficial effects of exercise on depression could be mediated via endorphins [57], monoamines [24], neutrophins [58], modifications of the hypothalamic–pituitary axis [59], or psychosocial mechanisms such as an increase of physical self-worth [60] and distraction [60,61]. Previous studies have rarely included measures of biological variables [62]. Disturbances in the hypothalamic–pituitary axis with high cortisol levels and the lack of ability to suppress endogen cortisol with dexamethasone [63] have long been known to accompany depression. Increased fitness has in experimental models shown to decrease cortisol response to psychological and physical stress [62,64]. Furthermore, serum prolactin has been used as an indicator of central serotonergic activity [25] which in one study has shown an abnormal response to physical activity in depressed patients [65].

In connection with the physical exercise testing at baseline and at the four-month follow-up we will measure cortisol and prolactin levels at rest and in response to exercise.

4.6. Physical activity

The exercise capacity is measured using bicycle ergometry as mentioned above. Physical activity during work and leisure time is measured at 4 and 12 months using a participant administered questionnaire by Saltin and Grimsby [66]. This questionnaire measures physical activity at work and leisure time separately, using four different categories.

5. Sample size calculation

With a Bonferoni adjusted alpha = 0.0166 (0.05 / 3) to allow comparison of the three arms of the trial [67], a minimal clinical relevant response on the HAM-D_{17} of 4 (δ), a standard deviation on the HAM-D_{17} of 6, $\beta = 0.1$, and a dropout of 40%, we calculated that we needed a total of 186 patients. However, entry data on the first 50 patients revealed a standard deviation of only 3 on the HAM-D_{17}, which was confirmed by the international literature [21,68–71]. We therefore adjusted our sample size calculation to a standard deviation on the HAM-D_{17} of 4. Based on this, 28 patients are required in each group, thus a total of 84 patients. With a 12-month dropout of 60% (based on the first 50 patients), we aim at including a total of 135 patients.

6. Statistics

The statistical analyses will be based on the intention-to-treat principle including all patients that have been randomized. For the primary outcome the mean change in HAM-D_{17} and change in lost days from work will be analyzed using ANOVA. In case of a significant difference among the groups, we will compare the non-aerobic with the aerobic training group. If the latter comparison shows no significant difference we will be combine them and compare them with the control group. In case of skewed attrition, in order to assess the influence of missing data on the 12-month symptom outcome, the data on the HAM-D_{17}, Becks Depression Inventory, and Montgomery Asbergs Depression Rating Scale will be subjected to further analysis in a repeated measurements model with unstructured variance matrix. This approach assumes that the distribution of missing data could be estimated from the information from previous interviews. The condition for using this method is the assumption that data are missing at random when taking into consideration the information extracted from baseline and four month follow-up interviews. Covariates entered in the repeated measurements model will be treatment, sex, and antidepressive medication. Accounting for the baseline value is automatic since it is included in the model, and no treatment effect is allowed for at baseline. An alternative approach to examine the influence of skewed attrition is sensitivity analysis, which also will be carried out.

Due to a possible effect of exercise on anxiety related items in the HAM-D_{17} [72,73] we will conduct an analysis of intervention effect using the HAM-D_{17} excluding the anxiety related items.

Per-protocol analysis will be carried out for patients having attended $\geq 50\%$ of their assigned intervention.

The two-sided P-value is set at ≤ 0.05. Due to the number of comparisons we employ Bonferoni adjusted P-value and we will interpret significant findings conservatively.

7. Ethical issues

The protocol has been accepted by the local ethics committee (KF 01-213) and registered at ClinicalTrials.gov (NCT 00103415). The results will be reported according to the CONSORT statement (www.CONSORT-Statement.org). The patients will be requested to give written informed consent after having received written and oral information on the trial following the Helsinki Declaration.

Only limited harm is expected to the patients in terms of muscle tenderness, venous puncture, etc. The staff instructioning the patients consist of highly qualified physiotherapists. In the event that one of the patients is getting worse or has an unresolved health problem, the physiotherapists are instructed to refer the patients to hospital or general practitioner depending on the nature of the problem.

8. Sponsor

The project is sponsored by ‘Forsikring og Pension’ (‘Assurance and Pension’), which is an organization promoting Danish insurance companies. The organization wanted ‘lost days from work’ to be included as an outcome measure.

Otherwise the sponsor has no role in trial design, data collection, data analysis, data interpretation, or in the future writing of the report. The corresponding author will have full access to all the data and will together with the co-authors make the decision when to submit results for publication.

9. Discussion

The strengths of this pragmatic trial are the method of randomization conducted by an independent institution providing adequate allocation concealment. Second, we intend to obtain blinded assessment of the primary outcome measure, HAM-D_{17}. Third, the intention-to-treat analysis gives this trial high internal validity. Fourth, the long follow-up time and the clinical relevant measures (HAM-D_{17}, lost working days) have the potential of reliably informing clinical practice. The realistic setting of the interventions makes the external validity of the trial strong. The referral from general practitioners to a four-month supervised exercise scheme is already a reality in the area, however, not for this particular group of patients. As mentioned above patients with depression have an increased all-cause mortality, which especially involves cardiovascular diseases, diabetes, hypertension, and possibly osteoporoses which all could be due to sedentary lifestyle. A sedentary lifestyle is increasingly recognized as a risk factor for a variety of diseases and this is indeed a part of the behavior of depressed patients.

Relaxation and massage exercises in the control group might have an anxiety reducing effect, thus influencing the HAM-D_{17} score in a positive manner and thereby increasing the risk of a type II error. However, if patients in the control group are not offered an attractive intervention we believe that attrition in the control group could bias the result of the trial. Therefore the activity in the control group must be perceived as relevant to the patients in order to keep them compliant. Due to the possible anxiety reducing effect of exercise and relaxation we plan an analyses on HAM-D_{17} after exclusion of anxiety related questions. This will enable us to examine the effect of the interventions on depression including the anxiety aspect and excluding the anxiety aspect. In order to reduce the confounding effect physiotherapist may have on patients on anxiety related aspects, we try to establish groups with equal number of patients per physiotherapist.

Based on the first 50 patients enrolled in the DEMO trial we know that a majority has received antidepressive medicine for more than 6 weeks. This could indicate that the patients are resistant to medical treatment and perhaps also to a possible positive effect of training which could lead to a type II error. Another limitation of this trial could be that patients are not blinded to the intervention, and only the interviewers and the biochemists measuring the biological values are blinded to intervention group.

Trials with multiple outcome measures or trials making more than one comparison face the problem of mass significance. It is in such trials difficult to define which p-value level that should be considered significant [74]. In the present DEMO trial we have tried to reduce the significance level due to the multiple comparisons we plan. Furthermore we will interpret significant findings conservatively.

During the last decade our knowledge has increased regarding the effect poorly conducted trials may have on estimates of intervention effects [19,20,75]. In this perspective the poor methodological quality of many trials within the field of exercise and depression should be interpreted with caution. Exercise advice from professionals to patients on this issue should be based on the same level of evidence as any medication given. The results from the DEMO trial will be published in 2008–2009.

References

[1] Andrews G, Henderson S, Hall W. Prevalence, comorbidity, disability and service utilisation: overview of the Australian National Mental Health Survey. Br J Psychiatry 2001;178:145–53.
[2] Blazer DG, Kessler RC, McGonagle KA, Swartz MS. The prevalence and distribution of major depression in a national community sample: the national comorbidity survey. Am J Psychiatry 1994;7:979–86.
[3] Spaner D, Bland RC, Newman SC. Major depressive disorder. Acta Psychiatr Scand 1994;Suppl 376:7–15.
[4] Tossani E, Cassano P, Fava M. Depression and renal disease. Sem Dial 2005;18:73–81.
[5] Lèpine JP, Gastpar M, Mendlewicz J, Tylee A. Depression in the community: the first pan-European study DEPRES (Depression Research in European Society). Int Clin Psychopharmacol 1997;12:19–29.
[6] Simon GE, Revicki D, Heiligenstein J, et al. Recovery from depression, work productivity, and health care costs among primary care patients. Gen Hosp Psych 2000;22:153–62.
[7] Mussolino ME, Jonas BS, Looker AC. Depression and bone mineral density in young adults: results from NHANES III. Psychosom Med 2004;66:533–7.

[8] Kahl KG, Rudolf S, Stoeckelhuber BM, et al. Bone mineral density, markers of bone turnover, and cytokines in young women with borderline personality disorder with and without comorbid major depressive disorder. Am J Psychiatry 2005;162:168–74.
[9] Meyer CM, Armenian HK, Eaton WW, Ford DE. Incident hypertension associated with depression in the Baltimore Epidemiologic Catchment area follow-up study. J Affect Disord 2004;83:127–33.
[10] Penninx BWJH, Rejeski WJ, Pandya J, et al. Exercise and depressive symptoms: a comparison of aerobic and resistance exercise effects on emotional and physical function in older persons with high and low depressive symptomatology. J Gerontol B Psychol Sci Soc Sci 2002;57:124–32.
[11] Ford DE, Mead LA, Chang PP, Cooper-Patrick L, Wang NY, Klag MJ. Depression is a risk factor for coronary artery disease in men: the precursors study. Arch Intern Med 1998;158:1422–6.
[12] Rugulies R. Depression as a predictor for coronary heart disease: a review and meta-analysis. Am J Prev Med 2002;23:51–61.
[13] Kessing LV, Andersen PK. Does the risk of developing dementia increase with the number of episodes in patients with depressive disorder and in patients with bipolar disorder? J Neurol Neurosurg Psychiatry 2004;75:1662–6.
[14] Ustun TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJL. Global burden of depressive disorders in the year 2000. Br J Psychiatry 2004;184:386–92.
[15] Greenberg PE, Kessler RC, Birnbaum HG, et al. The economic burden of depression in the United States: how did it change between 1990 and 2000. J Clin Psychiatry 2003;64(12):1465–75.
[16] 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;322:763–7.
[17] Brosse AL, Sheets ES, Lett HS, Blumenthal JA. Exercise and the treatment of clinical depression in adults: recent findings and future directions. Sports Med 2002;32:741–60.
[18] Kjaergaard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analysis. Ann Intern Med 2001;135:982–9.
[19] Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? The Lancet 1998;352:609–13.
[20] Juni P, Egger M. Allocation concealment in clinical trials. JAMA 2001;288:2407–8.
[21] Dunn AL, Trivedi MH, Kampert JB, Clark CG, Chambliss HO. Exercise treatment for depression: efficacy and dose response. Am J Prev Med 2005;28:1–8.
[22] Singh NA, Stavrinos TM, Scarbek Y, Galambos G, Liber C, Fiatarone Singh MA. A randomized controlled trial of high versus low intensity weight training versus general practitioner care for clinical depression in older adults. J Gerontol A Biol Sci Med Sci 2005;60:768–76.
[23] Martinsen EW, Hoffart A, Solberg Ø. Comparing aerobic with nonaerobic forms of exercise in the treatment of clinical depression: a randomized trial. Compr Psychiatry 1989;30(4):324–31.
[24] Chaouloff F. Effects of acute physical exercise on central serotonergic systems. Med Sci Sports Exerc 1997;29:58–62.
[25] Struder H, Weicker H. Physiology and pahtophysiology of the serotonergic system and its implications on mental and physical performance. Part II. Int J Sports Med 2001;22:482–97.
[26] Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–62.
[27] Bech P, Licht R, Stage K, et al. Rating scales for affektive lidelser. Hillerød: WHO Collaborating Centre; 2004.
[28] Andersen LB. A maximal cycle exercise protocol to predict maximal oxygen uptake. Scand J Med Sci Sports 1995;5:143–6.
[29] Bech P. Meta-analysis of placebo-controlled trials with mirtazapine using the core items of the Hamilton Depression Scale as evidence of a pure antidepressive effect in the short-term treatment of major depression. Int J Neuropsychopharmacol 2001;4:337–45.
[30] Kraemer WJ, Adams K, Cafarelli E, et al. Progression models in resistance training for healthy adults. Med Sci Sports Exerc 2002;34 (2):364–80.
[31] Gettman LR, Ayres JJ, Pollock ML, Jaffe AS. The effect of circuit weight training on strength, cardiorespiratory function, and body composition of adult men. Med Sci Sports Exerc 1978;10(3):171–6.
[32] Hass CJ, Feigenbaum MS, Franklin B. Prescription of resistance training for healthy populations. Sports Med 2001;31(14):953–64.
[33] Braith RW, Graves JE, Pollock ML, Leggett SL, Carpenter DM, Colvin AB. Comparison of 2 vs 3 days/week of variable resistance training during 10- and 18-week programs. Int J Sports Med 1989;10:450–4.
[34] Pollock ML, Gettman LR, Milesis CA, Bah MD, Durstine L, Johnson RB. Effects of frequency and duration of training on attrition and incidence of injury. Med Sci Sports Exerc 1977;9(1):31–6.
[35] Pollock ML, Gaesser GA, Butcher JD, et al. ACSM position stand: the recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. Med Sci Sports Exerc 1998;30:975–91.
[36] Pollock ML, Broida J, Kendrick Z, Miller HS, Janeway R, Linnerud AC. Effects of training two days per week at different intensities on middle-aged men. Med Sci Sports Exerc 1972;4(4):192–7.
[37] Atomi Y, Ito K, Iwasaki H, Miyashita M. Effects of intensity and frequency of training on aerobic work capacity of young females. J Sports Med 1978;18:3–9.
[38] Pollock ML, Miller HS, Linnerud AC, Cooper KH. Frequency of training as a determinant for improvement in cardiovascular function and body composition of middle-aged men. Arch Phys Med Rehabil 1975;56:141–5.
[39] Pedersen BK, Saltin B. Fysisk aktivitet-håndbog om forebyggelse og behandling. København: Sundhedsstyrelsen; 2003.
[40] Wenger HA, Bell GJ. The interactions of intensity, frequency and duration of exercise training in altering cardiorespiratory fitness. Sports Med 1986;3(5):346–56.
[41] Milesis CA, Pollock ML, Bah MD, Ayres J, Ward A, Linnerud AC. Effects of different durations of physical training on cardiorespiratory function, body composition, and serum lipids. The Res Q 1976;47(4):716–25.
[42] Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979;134:382–9.
[43] Beck AT, Steer RA, Brown GK. BDI-II manual. 2 ed. New York: Psychological Corporation; 1996.

[44] Hamilton M. Diagnoses and rating of anxiety. Br J Psychiatry 1969:76–9 [Speciel issue].
[45] Olsen LR, Mortensen EL, Bech P. The SCL-90 and SCL-90R versions validated by item response models in a Danish community sample. Acta Psychiatr Scand 2004;110:225–9.
[46] Vaynman S, Ying Z, Gomez-Pinilla F. Hippocampal BDNF mediates the efficacy of exercise on synaptic plasticity and cognition. Eur J Neurosci 2004;20:2580–90.
[47] Zheng H, Liu Y, Li W, et al. Beneficial effects of exercise and its molecular mechanisms on depression in rats. Behav Brain Res 2006;168:47–55.
[48] Ravnkilde B, Videbech P, Clemmensen K, Egander A, Rasmussen N, Rosenberg R. Cognitive deficits in major depression. Scand J Psychol 2002;43:239–51.
[49] Richards M, Hardy R, Wadsworth MEJ. Does active leisure protect cognition? Evidence from a national birth cohort. Soc Sci Med 2003;56:785–92.
[50] Nelson HE, Connel A. Dementia: the estimation of premorbid intelligence levels using the New Adult Reading Test. Cortex 1978;14:234–44.
[51] Wechscler D. The Wechsler adult intelligence scale—revised. San Antonio, Texas: The Psychological Corporation; 1981.
[52] Smith A. The serial sevens substraction test. Arch Neurol 1967;17:78–80.
[53] Reitan RM. The relation of the trail making test to organic brain damage. J Consult Psychol 1955;193:393–4.
[54] Borkowski JG, Benton AL, Spreen O. Word fluency and brain damage. Neuropsychologia 1967;5:135–40.
[55] Buschke H, Fuld PA. Evaluating storage, retension and retrievel in disordered memory and learning. Neurology 1974;24:1019–25.
[56] Meyers J, Meyers K. Rey complex figure and recognition trial: professional manual. Florida: Psychological Asessment Resources; 1995.
[57] Wildmann J, Kruger A, Schmole M, et al. Increase of circulating beta-endorphin-like imuunoreactivity correlates with the change in feeling of pleasantness after running. Life Sci 1986;38(11):997–1003.
[58] Mattson MP, Maudsley S, Martin B. A neural signaling triumvirate that influences ageing and age-related disease: insulin/IGF-1, BDNF and serotonin. Ageing Res Rev 2004;3:445–64.
[59] Dienstbier RA. Behavioral correlates of sympathoadrenal reactivity: the toughness model. Med Sci Sports Exerc 1991;23(7):846–52.
[60] McAuley E, Mihalko SL, Bane SM. Exercise and self-esteem in middle-aged adults: multidimensional relationships and physical fitness and self-efficacy influences. J Behav Med 1997;20:67–83.
[61] Nolen-Hoeksema S, Morrow J. Effects of rumination and distraction on naturally occurring depressed mood. Cogn Emot 1993;7:561–70.
[62] Nabkasorn C, Miyai N, Sootmongkol A, et al. Effects of physical exercise on depression, neuroendocrine stress hormones and physiological fitness in adolescent females with depressive symptoms. Eur J Public Health 2006;16(2):179–84 [cki159].
[63] Barden N. Implication of the hypothalamic–pituitary–adrenal axis in the physiophathology of depression. J Psychiatriy Neurosci 2004;29 (3):185–93.
[64] Traustadottir T, Bosch PR, Matt KS. The HPA axis response to stress in women: effects of aging and fitness. Psychoneuroendocrinology 2005;30:392–402.
[65] Kiive E, Maaroos J, Shlik J, Toru I, Harro J. Growth hormone, cortisol and prolactin responses to physical exercise: higher prolactin response in depressed patients. Pr0g Neuro-Psychopharmaco Biological Psychiatry 2004;28:1007–13.
[66] Saltin B, Grimsby G. Physiological analysis of middle-aged and old former athletes: comparison with still active athletes of the same age. Circulation 1968;38:1104–15.
[67] Leon AC. Multiplicity-adjusted sample size requirements: a strategy to maintain statistical power with Bonferroni adjustments. J Clin Psychiatry 2004;65:1511–4.
[68] Fava M, Alpert J, Nierenberg AA, et al. A double-blind, randomized trial of St John's wort, fluoxetine, and placebo in major depressive disorder. J Clin Psychopharmacol 2005;25:441–7.
[69] Lecrubier Y, Clerc G, Didi R, Kieser M. Efficacy of St. John's wort extract WS 5570 in major depression: a double-blind, placebo-controlled trial. Am J Psychiatry 2002;159:1361–6.
[70] Bjerkenstedt L, Edman GV, Alken RG, Mannel M. Hypericum extract LI 160 and fluoxetine in mild to moderate depression: a randomized, placebo-controlled multi-center study in outpatients. Eur Arch Psychiatry Clin Neurosci 2005;255:40–7.
[71] de Jonghe F, Hendricksen M, van Aalst G, et al. Psychotherapy alone and combined with pharmacotherapy in the treatment of depression. Br J Psychiatry 2004;185:37–45.
[72] Atlantis E, Chow CM, Kirby A, Fiatarone Singh M. An effective exercise-based intervention for improving mental health and quality of life measures: a randomized controlled trial. Prev Med 2004;39:424–34.
[73] Scully D, Kremer J, Meade MM, Graham R, Dudgeon K. Physical exercise and psychological well being: a critical review. Br J Sports Med 1998;32:111–20.
[74] Schulz KF, Grimes DA. Multiplicity in randomised trials I: endpoints and treatments. The Lancet 2005;365:1591–5.
[75] Kjaergaard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analysis. Ann Intern Med 2001;135:982–9.
[76] Olsen LR, Mortensen EL, Bech P. The SCL-92 and SCL-90R versions validated by item response models in a Danish community sample. Acta Psychiatr Scand 2004;110:225–9.
[77] Wechsler D. Manual for the Wechsler Adult Intelligence Scale (WAIS). San Antonio, Texas: The Psychological Corporation; 1955.