Comparing Physical Exercise in Groups to Group Cognitive Behaviour Therapy for the Treatment of Panic Disorder in a Randomized Controlled Trial

Anders Hovland

University of Bergen, and Solli District Psychiatric Centre (DPS), Nesttun, Norway

Inger Hilde Nordhus

University of Bergen, Norway

Trond Sjøbø, Bente A. Gjestad and Birthe Birknes

Solli District Psychiatric Centre (DPS), Nesttun, Norway

Egil W. Martinsen

University of Oslo, Norway

Torbjørn Torsheim and Ståle Pallesen

University of Bergen, Norway

Background: Previous studies have suggested that physical exercise can reduce symptoms for subjects suffering from panic disorder (PD). The efficacy of this intervention has so far not been compared to an established psychotherapy, such as cognitive behaviour therapy (CBT). Assessment of controlled long-term effects and the clinical significance of the treatment are also lacking. **Aim:** To compare physical exercise to CBT as treatment for PD, and assess controlled long-term and clinically significant effects. **Method:** PD-patients were randomized to either three weekly sessions of physical exercise ($n = 17$), or one weekly session of CBT ($n = 19$). Both treatments ran for 12 weeks, were manualized and administered in groups. Patients were assessed twice before the start of treatment, at post-treatment and at 6 and

---

Reprint requests to Anders Hovland, Department of Clinical Psychology, Christiesgt. 12, 5015 Bergen, Norway. E-mail: anders.hovland@psykp.uib.no An Appendix is also available online in the table of contents for this issue: http://journals.cambridge.org/jid_BCP

12 months thereafter. Primary outcome-measures consisted of the Mobility Inventory (MI), the Agoraphobia Cognitions Questionnaire (ACQ) and the Body Sensations Questionnaire (BSQ). **Results:** A two-way repeated measures MANOVA of these measures demonstrated a significant effect of time, $F(16, 544) = 7.28$, $p < .01$, as well as a significant interaction effect, $F(16, 544) = 1.71$, $p < .05$, in favour of CBT. This finding was supported by the assessment of clinically significant changes of avoidant behaviour and of treatment-seeking one year later. **Conclusion:** Group CBT is more effective than group physical exercise as treatment of panic disorder, both immediately following treatment and at follow-up assessments.

Introduction

Physical exercise has long been investigated as a treatment option for anxiety disorders. A large number of studies have demonstrated the ability of exercise to reduce anxiety levels, with the majority of these studies conducted using healthy volunteers (Long and van Stavel, 1995; Wipfli, Rethorst and Landers, 2008). Meta-analyses have also indicated that physical exercise can be effective for subjects with elevated levels of anxiety (Petruzzello, Landers, Hatfield, Kubitz and Salazar, 1991), and that it can reduce anxiety in patients suffering from somatic disorders (Herring, O'Connor and Dishman, 2010). However, few studies have evaluated the clinical efficacy of physical exercise among patients primarily suffering from an anxiety disorder. In one case report from 1974, Orwin used bouts of physical exercise to facilitate the treatment of situational phobia (Orwin, 1974). Sexton, Maere and Dahl (1989) found that physical exercise reduced anxiety levels in a mixed sample of anxious and depressed inpatients.

To the best of our knowledge, three studies have investigated the use of physical exercise as a treatment for patients with panic disorder (PD). Martinsen, Sandvik and Kolbjørnsrud (1989) conducted a naturalistic study with a mixed inpatient sample in which groups of eight patients participated. Exercise sessions were supervised with treatment involving at least one hour of exercise daily. A significant reduction of symptoms was observed among PD-patients following the exercise treatment. However, the treatment effect was not found to be significant at the 12-month follow-up assessment. Broocks et al. (1998) conducted a randomized controlled trial (RCT) comparing the effects on PD of physical exercise, placebo, or the tricyclic antidepressant clomipramine. Participants in that study were required to complete a four-mile route three times per week, either by walking or preferably by running. In addition, they participated in a supervised group exercise session once a week. The authors concluded that clomipramine produced a greater and more rapid treatment effect than physical exercise. However, physical exercise was comparable to clomipramine on certain outcome measures and was more efficacious than placebo. These findings were partly supported in Wedekind et al. (2010), where paroxetine and pill placebo were compared in combination with either exercise or relaxation. The physical exercise was conducted according to procedures in Broocks et al. (1998). The authors concluded that physical exercise was better than placebo, but less effective than paroxetine. Unfortunately, no follow-up assessment was conducted for either of these studies, leaving the controlled long-term effects of physical exercise largely unknown.

Physical exercise has also been found to effectively reduce anxiety sensitivity (Broman-Fulks, Berman, Rabian and Webster, 2004; Broman-Fulks and Storey, 2008; Smits et al.,

2008) in subjects with elevated levels of sensitivity to anxiety and anxiety symptoms. In these studies, physical exercise was performed for a total of six sessions over a 2-week period. Though these studies were conducted using subjects that were not diagnosed with PD, sensitivity to anxiety has been shown to be an important aspect of PD (Schmidt, Lerew and Jackson, 1999). Furthermore, it has been shown that physical exercise can have an acute anti-panic effect both in healthy volunteers (Esquivel, Schruers, Kuipers and Griez, 2002; Smits, Meuret, Zvolensky, Rosenfield and Seidel, 2009; Strohle et al., 2005) and in patients suffering from PD (Esquivel et al., 2008; Strohle et al., 2009). Which physiological and psychological mechanisms underlie the anxiolytic and anti-panic effects of physical exercise are still not clear. One possible mechanism is exposure to bodily sensations, which is known to be highly relevant to PD-patients (Chambless, Caputo, Bright and Gallagher, 1984).

In summary, an increasing amount of research suggests that physical exercise can be an important intervention for PD. However, to the best of our knowledge, no study has previously compared physical exercise to an established psychotherapy. Recent research has called for physical exercise to be compared to relevant treatment options, such as cognitive behaviour therapy (CBT; Broman-Fulks et al., 2004). Furthermore, it has been suggested that the clinical relevance of physical exercise, as well as its long-term effects on symptom improvement, should be assessed within a clinical context (Lawlor and Hopker, 2001; Salmon, 2001). CBT is currently regarded as the treatment of choice for PD (McHugh et al., 2007; National Institute for Health and Clinical Excellence, 2004), making it a suitable intervention against which physical exercise can be compared. It has shown that CBT for PD can be effectively delivered in group format (Telch et al., 1993).

The current study seeks to investigate the effects of physical exercise on PD in accordance with the recommendations mentioned above. By comparing group physical exercise to group CBT with a long-term follow-up, both the durability and the clinical effect of physical exercise can be assessed and compared with an established treatment option for PD. It was expected that physical exercise would yield large effects in its own right, though we predicted that between-groups comparisons would be in favour of CBT, based on results from previous studies (Clark et al., 1994; Broocks et al., 1998). Furthermore we wanted to assess the magnitude of long-term, clinically significant changes of group physical exercise compared to group CBT.

Method

Participants

Thirty-six individuals between the ages of 18 and 50 years met the diagnostic criteria for PD with or without agoraphobia according to the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR; American Psychiatric Association, 2000). The presence of PD was established through the administration of the Structured Clinical Interview for DSM-IV axis I Disorders (SCID-I; First, Spitzer, Gibbon and Williams, 1995) by a trained clinical psychologist (the first author). Participants were excluded if they fulfilled any of the following criteria at the time of the study: (1) brain-organic disorders; (2) psychotic disorders; (3) substance-abuse, including habitual use of benzodiazepines; (4) medical conditions that precluded participation in physical exercise; or (5) severe major depressive episode.

Exclusion according to criteria (1), (3) and (5) was established through the administration of the SCID-I, whereas criteria (2) and (4) were assessed by interview with the participant, and if necessary, by a consultation with the participant's general practitioner (GP). All participants were also interviewed with the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II; First, Spitzer, Gibbon, Williams and Benjamin, 1994) prior to treatment. For health and safety reasons, all participants were screened for risk-factors related to heart disease (e.g. a family history of cardiac disease, hypercholesterolemia, diabetes, hypertension or smoking) in co-operation with a cardiologist. Upon detection of risk-factors, the participant's GP was contacted with a request for a cardiac evaluation. Participants using psychotropic medication were requested to stabilize this use both during the treatment period and for 2 weeks before and after this period. Their use of psychotropic medication was monitored throughout participation.

Recruitment

The study was conducted in a day-care unit at a district psychiatric centre (DPS) in the municipality of Bergen, Norway. All GPs, psychologists, psychiatrists and primary health care services in the region (a geographic area with a population of 45,000) were invited to refer patients suffering from PD for treatment. The majority of participants were, however, recruited through advertisements in the local press (83%).

All participants were required to provide written informed consent prior to enrolment. The present study was approved by the Regional Committee for Medical and Health Research Ethics in Western Norway and by the Norwegian Social Science Data Services. The study is registered at ClinicalTrials.gov (identifier: NCT01076777).

Procedure

Initial contact was established by telephone, where subjects were provided with information about the nature of the study. At this stage, potential participants were screened for eligibility according to the inclusion and exclusion criteria by a clinical psychologist (the first author). Subjects who potentially fulfilled the criteria for participation were scheduled for an on-site evaluation, during which informed consent was obtained. Subjects were prior to treatment initiation informed that both interventions had been shown to be effective for panic disorder, but that these effects were better documented for CBT. Following the signing of the informed consent a structured clinical interview (SCID-I) was conducted. Once eligibility was established, participants completed the baseline assessment. As both interventions were conducted in groups of eight or less participants, a sufficient number (twice the maximal group, e.g. 16) of participants had to be recruited before randomization could take place. Therefore, depending on when they entered the study, participants had to wait a variable period (a mean of 68 days) following the baseline assessment before the pre-treatment assessment was conducted. Participants were then immediately randomized and allocated to one of the two interventions. Further assessments were conducted immediately post-treatment and at 6- and 12-month follow-ups. With the exception of the 12-month follow-up, all assessments were conducted on-site at the DPS. For the 12-month follow-up assessment, the outcome measures were sent by postal mail

to the participants. A total of three rounds of recruitments, allocations and treatments were completed.

Allocation.   Participants were randomly assigned to either of the two treatment conditions. Randomization was performed electronically (http://www.randomizer.org) at a location removed from the treatment centre. The person (the last author) conducting the randomization, received only reference-codes, and was hence completely blind to the participants' identities and personal information. A matched pairwise randomization procedure was used with pairs based on sex and on the Beck Anxiety Inventory score (BAI; Beck and Steer, 1993) recorded at the pre-treatment assessment.

Power

The effect size estimates were based on single studies on physical exercise and CBT with comparable participants and methods, as well as comparable or equivalent primary outcome measures (Broocks et al., 1998; Clark et al., 1994; Clark et al., 1999). An effect size (Cohen's *d*) of at least 0.9 in favour of CBT was expected when comparing the interventions on primary outcome measures. With an allocation ratio of 1, an alpha level of .05, and a power-level of .80 (one-tailed), analysis using Gpower (Faul, Erdfelder, Lang and Buchner, 2007) indicated that a total number of 32 subjects (16 in each group) would be sufficient.

Therapists

Physical exercise.   All sessions of physical exercise were led by a team of three treatment providers: one specialized physiotherapist, one occupational therapist and one psychiatric nurse. Both the physiotherapist and the psychiatric nurse had extensive experience (>10 years) in supervising physical exercise for groups of psychiatric patients. The occupational therapist had top level track and field experience and was highly experienced in coaching and supervising athletes.

CBT.   The cognitive behaviour therapy was conducted by a team of two specialist clinical psychologists with post-graduate qualifications in cognitive therapy. The two therapists had more than 15 years of experience as clinical psychologists, including more than 5 years of experience with the clinical application of cognitive therapy. Both had more than 5 years of experience in administering group treatment.

Interventions

Both treatments were administered in groups and ran for 12 weeks. Physical exercise was conducted three times per week, and CBT was conducted once a week. For each intervention, a booster session was offered 3 months following the termination of the treatment period. There was no contact between participants across interventions. An overview and comparison of the interventions is provided in the Appendix.

Physical exercise.   Physical exercise was conducted according to a manual designed at the clinic by the physiotherapist from the treatment team and the first author. This manual was developed in cooperation with an expert in the field of physical exercise (EWM). The

frequency and duration of exercises were based on recommendations from previous research on physical exercise with subjects suffering from anxiety (Martinsen, Hoffart and Solberg, 1989; Meyer and Broocks, 2000; Petruzzello et al., 1991) and recommendations regarding physical exercise for healthy adults (Haskell et al., 2007). Content in each session was also in accordance with the specific recommendations for these types of exercise (Pollock et al., 1998). As recommended (Meyer and Broocks, 2000), initial treatment sessions progressed carefully to allow the participants to adjust to the treatment stimulus. Studies have shown that, compared with healthy volunteers, anxiety symptoms increase as a function of exercise intensity for PD-patients, but that when they manage to perform the exercise, the likelihood of anxiety and panic is reduced (Stein et al., 1992). A pilot including eight patients previously admitted at the centre was conducted to ensure the practical feasibility and motivational properties of the manual. The manual consisted of three different types of sessions that were repeated weekly on set days. Day 1 focused on increasing aerobic fitness, and sessions consisted of 60 minutes of long-distance walking/running outdoors in steep terrain with interval exercises performed at the end of sessions. Intensity was monitored through the use of fitness heart rate monitors (Polar F4^{TM} Fitness Heart Rate Monitor with chest strap), and participants were instructed to stay within the pre-programmed zone of 60–80% of maximum heart rate (HR$_{max}$) for two-thirds of the session. HR$_{max}$ was calculated by the fitness heart rate monitor according to the formula HR$_{max}$ = 220 – age; the standard formula for estimating HR$_{max}$ based on such devices (Nes, Janszky, Wisløff, Støylen and Karlsen, 2012). Adherence to the prescribed intensity was monitored by the supervisors.

Day 2 focused on increasing muscular strength through circuit training. The circuit consisted of 9 different exercises that were performed for 60 seconds, with a 15-second break between each exercise and a 2-minute break between each circuit. Session durations increased from 30 to 45 minutes as the number of circuits correspondingly increased from 2 to 3 as treatment progressed. On day 3, exercises that varied in intensity were performed, including sports and games with elements of competition. These exercises lasted approximately 60 minutes. All sessions were preceded by an introduction, warm-up exercises and stretching, and ended with a short debriefing. Due to this the total duration for each session was approximately 90 minutes. The manual was designed to ensure that the duration and intensity of the various exercises increased over the treatment period. The participants were also supervised individually to ensure compliance and progression, and intensities were generally high across the different types of sessions.

Participants were initially briefed regarding the physical sensations they might experience during exercises and the similarity of these sensations to certain anxiety symptoms. This was done in accordance with previous knowledge about the initiation of physical exercise in this group of patients (Meyer and Broocks, 2000). During treatment sessions, conversation focused on the execution of the different exercises and practical issues pertaining to exercise in general. Treatment was always administered by two members of the treatment team. Session attendance was monitored, and participants who were unable to participate on a given occasion were instructed to perform an equivalent exercise on their own and to record details of the exercise on a provided form (this was not recorded as treatment attendance).

CBT.   Therapy was based on the model for PD developed by David Clark and colleagues, with an emphasis on cognitive restructuring and behavioural experiments (Clark, 1986; Clark et al., 1994, 1999). Although Clark et al. refer to this treatment as cognitive therapy, treatment

in the current study included other components that are central to CBT for PD (Gould, Otto and Pollack, 1995), including exposure to somatic sensations as well as situational exposure. Therapy is therefore referred to as CBT. The treatment manual was developed by the therapists, with the content and treatment progression customized for group sessions. The manual was evaluated by two independent experts on cognitive treatment of PD (Asle Hoffart, personal communication, 28 September 2007; Hans M. Nordahl, personal communication, 28 August 2007), and was judged to be in accordance with the panic model by David Clark and colleagues. Treatment content focused on identifying catastrophic misinterpretations, assessing safety behaviours and establishing alternative hypotheses for the bodily sensations. Exposure to somatic sensations was performed in group using hyperventilation and physical exertion. Situational exposure was mainly performed individually as homework between sessions. When deemed necessary for further progress, participants performed situational exposure accompanied by one of the therapists during part of the final sessions. Treatment was always administered by both members of the treatment team. Every session lasted approximately 2 hours, and included a brief assessment of symptom severity, as well as evaluation and planning of homework.

Treatment fidelity

As CBT represented the control condition in the present study, additional measures were taken to ensure that the quality and integrity of this intervention were satisfactory. A total of six treatment-sessions were randomly selected, evaluated and rated, thus covering half of the total treatment content for the intervention. This evaluation was performed on-site by an observer (a specialist clinical psychologist that was formally qualified as a supervisor of cognitive therapy by the Norwegian Association for Cognitive Therapy). The observer had personal experience with delivering CBT for PD. The quality of the cognitive therapy was scored on the Cognitive Therapy Adherence and Competence Scale (CTACS; Barber, Liese and Abrams, 2003). The random selection of sessions was devised to ensure that the rated sessions were selected both from the initial as well as the final stages of the treatment period. Between sessions, the therapists received supervision from a different specialist in clinical psychology who also was formally qualified as a supervisor of cognitive therapy.

Measures

All outcome measures were adapted from English to the participant's native language by a translation and back-translation procedure.

   Primary outcome measures: Agoraphobia.   The Agoraphobia Cognitions Questionnaire (ACQ; Chambless et al., 1984) and the Mobility Inventory (MI; Chambless, Caputo, Jasin, Gracely and Williams, 1985) were used to evaluate agoraphobic cognitions and behaviours respectively. The ACQ is a 14-item questionnaire assessing the frequency of thoughts that are common in agoraphobia and is considered to be as measure of fear of fear. Items are rated on a scale ranging from 1 (thought never occurs when I am nervous) to 5 (thought always occurs when I am nervous). The Cronbach's alpha for ACQ at baseline in the present study was .75. The MI is a 27-item questionnaire designed to assess the frequency of avoidance for a range of situations relevant to daily life. The MI consists of two subscales assessing avoidance for

these situations both when alone and when accompanied. Items are rated on a scale ranging from 1 (never avoid) to 5 (always avoid). The subscales of MI are here referred to as MI-Alone and MI-Accompanied. The two MI subscales also provide estimates of an individual's actual freedom of movement, and therein their level of recovery through treatment. In the present study, Cronbach's alphas at baseline for MI-Alone and MI-Accompanied were .91 and .88, respectively.

Bodily sensitivity.   Alertness to bodily sensations was measured by the Body Sensations Questionnaire (BSQ; Chambless et al., 1984). The BSQ is a 17-item questionnaire designed to measure fear of bodily sensations, and it provides a list of sensations that may occur when the person is anxious or in a feared situation. Items appear on a scale ranging from 1 (not at all frightened by this sensation) to 5 (extremely frightened by this sensation). The Cronbach's alpha at baseline for the BSQ in the present study was .87.

Panic Attack Scale.   Both panic frequency and panic-related distress and disability were assessed together using a 2-item scale, with one item for each domain. The scale was completed by the participants as well as by a clinical psychologist (the first author). This corresponds to the procedure described by Clark et al. (1994, 1999). Panic attack frequency was measured on a 5-point scale ranging from 0 (no panic attacks) to 4 (one or more panic attacks per day). Panic-related distress and disability was measured on a 9-point scale running from 0 (not at all disturbing) to 8 (very disturbing).

Secondary outcome measures: General anxiety.   General anxiety was measured both with the Beck Anxiety Inventory (BAI; Beck and Steer, 1993) and the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, Lushene, Vagg and Jacobs, 1983). The BAI is comprised of 21 items describing subjective, somatic, or panic-related symptoms. Each item is scored on a scale with values ranging from 0 (not at all) to 3 (severely – I could barely stand it). The Cronbach's alpha at baseline for the BAI in the present study was .89. The STAI consists of two separate 20-item scales designed to assess both state and trait levels of anxiety. Both scales ask subjects to rate the degree to which each item applies to them on a scale ranging from 1 (not at all) to 4 (very much). The STAI-S (state) is comprised of a list of 20 statements that assess the degree of anxiety and tension one feels at the current moment. The STAI-T (trait) is comprised of a list of 20 different statements regarding the degree of anxiety and tension one feels in general. In the present study, Cronbach's alpha at baseline for both the STAI-S and the STAI-T was .91.

Depression.   Level of depression was assessed using the Beck Depression Inventory II (BDI-II; Beck, Steer and Brown, 1996). The inventory is comprised of 21 items, all of which were scored on a 4-point scale (0–3) that aimed to reflect the levels of severity of the symptom in question. The Cronbach's alpha at baseline for the BDI-II in this study was .87.

Quality of life.   The Quality of Life Inventory (QoLI; Frisch, Cornell, Villanueva and Retzlaff, 1992) assesses 16 different areas of life. Each area is rated on two scales, with one scale measuring the importance of this particular area to the individual, and the other the individual's level of satisfaction with the area in question. The score for each area is made up by the product of the scores for satisfaction (rated on a 6-point scale ranging from -3 to 3) and importance (rated on a 3-point scale ranging from 0 to 2). Areas rated with a zero for

importance were not included in the calculation of the overall mean (ranging from -6 to 6). The Cronbach's alpha at baseline for the QoLI in this study was .80.

Additional measures: Quality of therapy. The Cognitive Therapy Adherence and Competence Scale (CTACS; Barber et al., 2003) is a 21-item rating scale assessing the quality of cognitive therapy that covers both adherence to the therapy manual and the competence with which therapy is delivered. Items cover various aspects deemed relevant to the adequate delivery of cognitive therapy, with each item rated from 0 (poor) to 6 (excellent) for both adherence and competence. Based on this, a mean score may be calculated for each domain. In addition, one item specifically assesses the therapist's overall performance. The Cronbach's alpha over the 6 ratings in this study was .93 for both the adherence-items and competence-items.

Six-minute walk test. The 6-minute walk test (6MWT) is a self-paced measurement of sub-maximal levels of functional capacity. It is a simple test in which subjects are instructed to walk, and not run, as far as possible within 6 minutes on a flat and straight corridor. It has been widely used in samples where participants suffer from cardiac or pulmonary disease, but also with subjects with other impairments, such as fibromyalgia, and has been shown to be both valid and reliable in such samples (King et al., 1999). Assessment in the current study was conducted according to developed guidelines (Crapo et al., 2002), with a corridor of 12.5 meters.

All of the above measures except the 6MWT were administered five times: at baseline, pre-treatment, post-treatment, the 6-month follow-up and the 12-month follow-up. The 6MWT was assessed only at pre- and post-treatment. As the 12-month follow-up assessment was conducted via postal mail, the clinician rated versions of panic frequency and panic-related distress and disability were not administered at this assessment. For the 12-month follow-up assessment, two additional questions were provided in addition to the standard questionnaires referred to above. These two questions assessed treatment-seeking behaviour as well as changes in medication that occurred during the year following the treatment period and were phrased in the following manner (here translated into English): 1) "Have you entered into a new therapeutic treatment directed at your panic disorder following termination of the treatment in the study?" and 2) "Have you been prescribed new psychotropic medication directed at your panic disorder, or changed your current prescription, following the termination of the treatment in the study?". For both questions, participants were prompted to provide further information if they responded positively.

Statistical procedures

Differences in the characteristics of the participants assigned to the two interventions were assessed using *t*-tests for independent samples for the continuous variables and the Fisher's exact test for the categorical variables.

All subjects were included in all analyses of outcome according to the principle of intention-to-treat (ITT). Last Observation Carried Forward (LOCF) was used for the participants who failed to complete treatment, declined to participate at either one or both follow-up assessments, or where contact could not be established. For the inventories, missing data were replaced by the mean of the recorded values, provided that no more than 20% of

the values were missing. If more than 20% of the values were missing, the data were deemed invalid, and values were substituted by LOCF.

Due to the multiple outcome measures in this study, analyses were performed in a two-step procedure in order to avoid loss of power due to multiple testing and to reduce the likelihood of a type I error. Initially, four doubly two-way (Time×Group) repeated measure MANOVAs were performed in order to investigate the main effects on the primary and secondary outcome measures. For the primary outcome measures, the first MANOVA included the following four inventories: MI-Alone, MI-Accompanied, ACQ and BSQ. Separate MANOVAs were performed on the single item scales of panic frequency and panic-related distress and disability. As the clinician-rated versions were not administered at the 12-month follow-up, separate MANOVAs had to be conducted for the self-report versions and the clinician-rated version of these scales. Finally, all secondary outcome measures were analyzed together in one MANOVA. If a significant Time×Group effect was detected, separate two-way (Time×Group) repeated measures ANOVAs were performed on the implicated outcome measures. Within-groups effects of treatment with reference to baseline levels were assessed using *t*-tests, and bias from multiple comparisons was controlled for with Bonferroni-corrections. All analyses were conducted using PASW (SPSS) version 17.0.

Effect sizes.   Effect-sizes (ES) are reported as Cohen's *d*, based on pooled standard deviations (Cohen, 1988).

Reliable and clinically significant change.   Reliable and clinically significant changes were analyzed according to the recommendations by Jacobson and Truax (1991). The Mobility Inventory–Alone subscale was considered the most suitable for assessing clinical improvement and recovery, as it aims to measure an individual's freedom (or restriction) of movement. For change within a subject to be considered reliable, it had to exceed the calculated Reliable Change Index (RCI). For the assessment of clinically significant change, method C as referred to in Jacobson and Truax (1991) was considered the most appropriate. According to this method, a midpoint between the normal and the clinical population is estimated. When a subject's mean score falls closer to the mean of the normal population than to that of the clinical population, the change is considered to be clinically significant, provided that this change is also reliable according to the RCI. For a subject to be considered recovered, he or she also had to fulfil an additional criterion as recommended by D. Clark (personal communication, 22 June 2009) of having zero panic attacks during the last 2 weeks prior to assessment.

Results

Participant flow and attrition

Of the 141 subjects who made contact with the study's staff, 66 were invited for a detailed diagnostic interview (see Figure 1 for an overview of participants' flow in the study). The most frequent reasons for exclusion at this stage were either that the symptoms did not indicate the presence of PD, or that subjects declined participation when informed about the nature of the study. Of the 66 invited subjects, 36 participants were randomized to either of the two treatment conditions. Excluded subjects either failed to match inclusion and exclusion criteria, or declined to participate. A total of four participants (11% of the sample) were
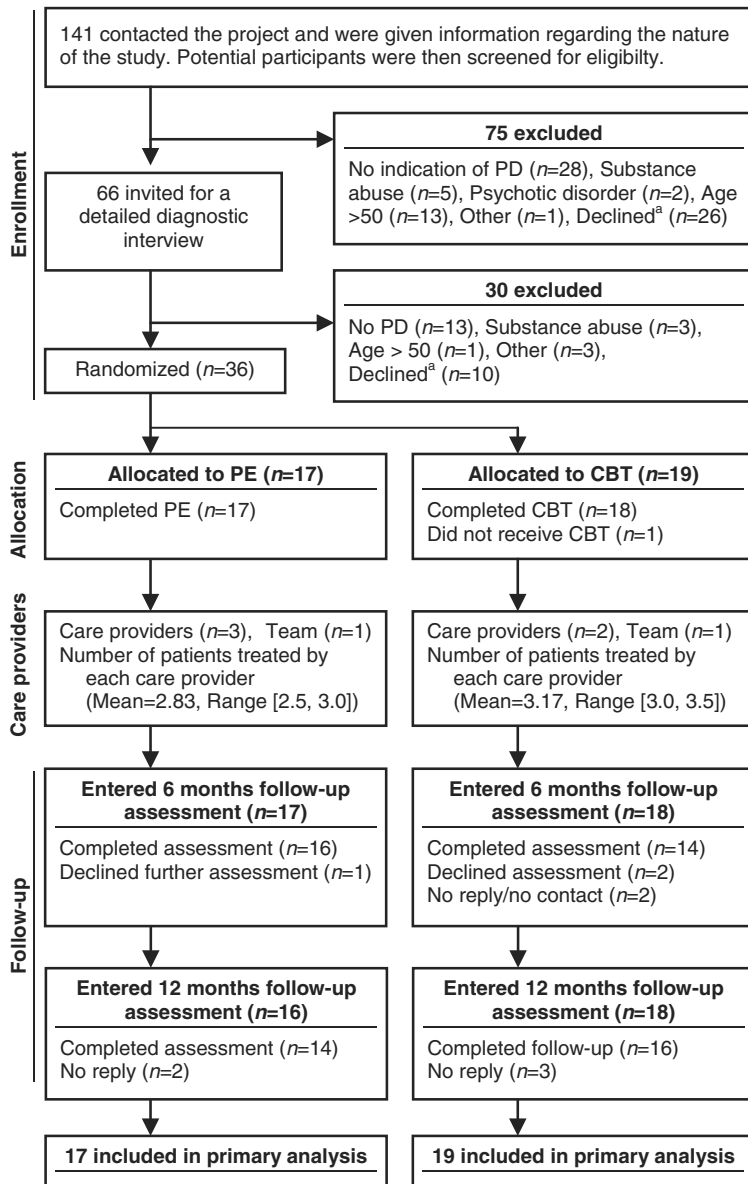
Figure 1. Consort diagram displaying the process of inclusion, allocation and attrition. PE = physical exercise

[a] Includes subjects whose job, studies or distance of travel prevented participation.

referred for a cardiac evaluation prior to treatment. Only one participant dropped out during the course of treatment (in the CBT-intervention following the second treatment session). Thus, 35 participants received treatment as planned. For various reasons, some participants did not complete the follow-up assessments (see Figure 1). At the 6-month follow-up, one participant from the physical exercise group declined assessment, while a total of four from the CBT group were not assessed. At the 12-month follow-up, a total of three participants in each condition were not assessed. Viewed together, only three participants were not assessed at either one of the follow-up assessments: one in the physical exercise-condition and two in the CBT-condition.

Baseline characteristics

Table 1 presents the demographic and clinical characteristics of the participants at baseline. This Table shows that 80.6% of the participants were female, that 27.7% reported living alone, and that 47.2% reported full-time employment. Participants had a mean age of 37.9 years, and a mean of 13.6 total years of education. The mean duration of PD was 10.1 years, and 38.9% of the sample had previously sought psychotherapy for PD or PD-related symptoms. Tests of physical fitness indicated no differences between the groups, neither according to the 6-minute walk test, nor in terms of on-going levels of physical exercise at baseline. The sample had a mean Body Mass Index of 26.7. Of the total sample, 36.1% used a selective serotonin reuptake inhibitor (SSRI), and 13.9% reported to have used benzodiazepines intermittently within the last month. As expected, a large proportion (80.6%) of the participants was also diagnosed with agoraphobia, and depression was found to be present in 38.9% of the sample. On average, participants were diagnosed with a mean of 2.1 comorbid axis 1 disorders and 0.4 comorbid axis 2 disorders. The distribution of the various comorbid axis 1 disorders is presented in detail in Table 1. There were no significant differences between the groups in the listed characteristics. Table 2 summarizes the results for both groups on all outcome measures and shows that there were no significant differences between the groups prior to treatment. Furthermore, the Bonferroni-corrected within-groups comparisons did not indicate any significant changes from baseline to pre-treatment.

Treatment attendance

The physical exercise group had a mean attendance rate of 88.7% ($SD = 10.0$) for treatment sessions. The corresponding rate for the CBT group was 86.0% ($SD = 23.6$).

Psychotropic medication

Each subject's use of psychotropic medication was monitored throughout participation, and converted to daily doses. One-way repeated measures ANOVAs including all five assessments were conducted for each intervention, indicating that levels of medication remained stable for participants in both the physical exercise group, $F(2.59, 41.41) = 1.58, p = .213$ (Huynh-Feldt correction) and the CBT group, $F(1.70, 30.55) = 0.48, p = .593$ (Huynh-Feldt correction).

Table 1. Baseline characteristics of patients assigned to Physical Exercise (PE) or CBT

Characteristic	All patients (N = 36)		PE (N = 17)		CBT (N = 19)	
	N	%	N	%	N	%
Female	29	80.6	15	88.2	14	73.7
Marital status (e.g. living alone)	10	27.8	7	41.2	3	15.8
Employment 6 months prior to treatment						
Full-time employed	17	47.2	10	58.8	7	36.8
Part-time employed	10	27.8	4	23.5	6	31.6
In rehabilitation or on sick leave	4	11.1	1	5.9	3	15.8
Without regular work	3	8.3	2	11.8	1	5.3
Student	2	5.6	0	0	2	10.5
	Mean	SD	Mean	SD	Mean	SD
Age (years)	37.9	8.6	38.1	8.6	37.8	8.9
Years of education	13.6	2.5	13.9	2.5	13.4	2.5
Duration of PD	10.1	9.5	12.3	10.7	8.2	8.1
Test of physical fitness level (6MWT)	570.5	90.1	593.1	70.5	550.3	103.7
Body Mass Index	26.7	5.8	26.9	6.2	26.6	5.5
	N	%	N	%	N	%
Physical exercise[a]	10	27.8	5	29.4	5	26.3
Medication						
SSRIs	13	36.1	7	41.2	6	31.6
Benzodiazepines[b]	5	13.9	3	17.6	2	10.5
Previous treatment for PD	14	38.9	6	35.3	8	42.1
Agoraphobia	29	80.6	13	76.5	16	84.2
Axis 1 comorbidity						
Depression	14	38.9	6	35.3	8	42.1
Social phobia	15	41.7	8	47.1	7	36.8
Single phobia	17	47.2	7	41.2	10	52.6
Generalised Anxiety Disorder	16	44.4	8	47.1	8	42.1
Health anxiety	8	22.2	5	29.4	3	15.8
Obsessive-Compulsive Disorder	3	8.3	1	5.9	2	10.5
	Mean	SD	Mean	SD	Mean	SD
Number of comorbid axis 1 disorders	2.1	1.2	2.2	1.4	2.1	0.9
Number of comorbid axis 2 disorders	0.4	0.6	0.4	0.6	0.4	0.6

Notes: No significant ($p < .05$) group differences were found. Both 6MWT and Body Mass Index (BMI) were assessed at pre-treatment. [a]Refers to ongoing (within last 14 days) level of physical exercise recommended for cardiorespiratory fitness or muscular strength as recommended by ACSM (Pollock et al., 1998).[b]Refers to the intermittent use of benzodiazepines (Valium, Stesolid, Imovane) within the last month.

Analysis of outcome measures

    Between-groups effects: primary outcome measures. Please refer to Table 2 for values, effects, and the comparisons of the two treatments. Doubly two-way (Time×Group) repeated

Table 2. Analyses of outcome measures

	Treatments				Comparisons		
	Physical Exercise (n = 17)		CBT (n = 19)		Between-groups effects		Time × Group
Outcome measures	Mean (*SD*)	ES	Mean (*SD*)	ES	ES	95% CI	F
Primary Outcome Measures – Inventories – MANOVA							1.71 (16, 544)*
Mobility Inventory (MI) – Alone – ANOVA							3.75 (2.28, 77.61)*
Baseline	2.42 (0.57)		2.37 (0.78)		0.07	[−0.58, 0.73]	
Pre-treatment	2.38 (0.65)	−0.06	2.40 (0.70)	0.04	−0.02	[−0.68, 0.63]	
Post-treatment	1.98 (0.74) [a]	−0.65	1.55 (0.55) [a]	−1.18	0.65	[−0.02, 1.32]	
6-month follow-up	1.93 (0.75) [a]	−0.71	1.59 (0.76) [a]	−0.98	0.44	[−0.22, 1.10]	
12-month follow-up	2.04 (0.68) [a]	−0.59	1.61 (0.75) [a]	−0.98	0.59	[−0.08, 1.26]	
Mobility Inventory (MI) – Accompanied – ANOVA							3.93 (2.19, 74.28)*
Baseline	1.79 (0.43)		1.98 (0.70)		−0.32	[−0.97, 0.34]	
Pre-treatment	1.79 (0.55)	0.01	1.92 (0.58)	−0.09	−0.21	[−0.87, 0.45]	
Post-treatment	1.61 (0.61)	−0.33	1.39 (0.53) [a]	−0.92	0.36	[−0.29, 1.02]	
6-month follow-up	1.51 (0.55)	−0.54	1.44 (0.62) [a]	−0.79	0.12	[−0.53, 0.78]	
12-month follow-up	1.64 (0.54)	−0.30	1.41 (0.58) [a]	−0.87	0.41	[−0.25, 1.07]	
Agoraphobic Cognitions Questionnaire (ACQ) – ANOVA							1.15 (2.11, 71.62)
Baseline	2.30 (0.52)		2.23 (0.61)		0.13	[−0.52, 0.79]	
Pre-treatment	2.23 (0.45)	−0.16	2.12 (0.49)	−0.20	0.23	[−0.43, 0.89]	
Post-treatment	1.85 (0.58) [a]	−0.80	1.50 (0.38) [a]	−1.40	0.70	[0.03, 1.38]	
6-month follow-up	1.75 (0.52) [a]	−1.04	1.44 (0.33) [a]	−1.58	0.72	[0.05, 1.40]	
12-month follow-up	1.76 (0.55) [a]	−0.99	1.50 (0.35) [a]	−1.45	0.57	[−0.10, 1.24]	
Body Sensations Questionnaire (BSQ) – ANOVA							4.08 (2.81, 95.43)*
Baseline	2.86 (0.60)		2.72 (0.75)		0.20	[−0.45, 0.86]	
Pre-treatment	2.55 (0.55)	−0.54	2.48 (0.71)	−0.32	0.10	[−0.56, 0.75]	
Post-treatment	2.25 (0.80) [a]	−0.85	1.55 (0.43) [a]	−1.87	1.07	[0.37, 1.77]	
6-month follow-up	2.13 (0.68) [a]	−1.11	1.57 (0.54) [a]	−1.73	0.90	[0.21, 1.59]	
12-month follow-up	2.12 (0.69) [a]	−1.12	1.70 (0.63) [a]	−1.45	0.64	[−0.03, 1.31]	
Primary Outcome Measures – Single Item Scales – Self-report – MANOVA							1.00 (8, 272)
Panic-related distress/disability (0–8; Self-report)							
Baseline	5.94 (2.05)		6.42 (1.64)		−0.25	[−0.91, 0.40]	
Pre-treatment	5.76 (1.75)	−0.09	6.11 (1.49)	−0.19	−0.21	[−0.87, 0.44]	
Post-treatment	2.82 (2.74) [a]	−1.26	1.79 (1.87) [a]	−2.57	0.43	[−0.23, 1.10]	
6-month follow-up	2.00 (2.24) [a]	−1.80	1.74 (2.05) [a]	−2.47	0.12	[−0.54, 0.77]	
12-month follow-up	3.06 (2.88) [a]	−1.13	2.00 (2.13) [a]	−2.27	0.41	[−0.25, 1.07]	
Panic frequency (0–4; Self-report)							
Baseline	1.47 (1.13)		1.79 (1.32)		−0.25	[−0.91, 0.40]	
Pre-treatment	1.35 (1.12)	−0.10	1.68 (1.34)	−0.08	−0.26	[−0.92, 0.40]	
Post-treatment	0.76 (1.25)	−0.58	0.68 (0.95)	−0.95	0.07	[−0.58, 0.73]	
6-month follow-up	0.53 (1.13)	−0.82	0.58 (0.90) [a]	−1.05	−0.05	[−0.70, 0.61]	
12-month follow-up	0.82 (1.13)	−0.56	0.47 (0.77) [a]	−1.20	0.36	[−0.30, 1.02]	
Primary Outcome Measures – Single Item Scales – Clinician-rated – MANOVA							0.75 (6, 204)
Panic-related distress/disability (0–8; Clinician-rated)							
Baseline	6.53 (1.42)		6.79 (1.32)		−0.19	[−0.84, 0.47]	
Pre-treatment	6.24 (1.72)	−0.18	6.21 (1.81)	−0.36	0.02	[−0.64, 0.67]	
Post-treatment	2.82 (2.79) [a]	−1.64	2.37 (2.50) [a]	−2.17	0.17	[−0.49, 0.82]	
6-month follow-up	2.59 (2.81) [a]	−1.73	2.11 (2.31) [a]	−2.44	0.18	[−0.47, 0.84]	
Panic frequency (0–4; Clinician-rated)							
Baseline	1.35 (1.12)		1.79 (1.36)		−0.34	[−1.00, 0.31]	
Pre-treatment	2.00 (1.32)	0.52	1.58 (1.39)	−0.15	0.30	[−0.36, 0.96]	
Post-treatment	0.76 (1.25)	−0.49	0.63 (0.90) [a]	−0.99	0.12	[−0.54, 0.77]	
6-month follow-up	0.65 (1.22)	−0.58	0.58 (0.96) [a]	−1.01	0.06	[−0.59, 0.72]	

Table 2. Continued.

Outcome measures	Physical Exercise (n = 17)		CBT (n = 19)		Between-groups effects		Time × Group
	Mean (SD)	ES	Mean (SD)	ES	ES	95% CI	F
Secondary Outcome Measures – Inventories – MANOVA							1.16 (20, 540)
Beck Anxiety Inventory (BAI)							
Baseline	27.76 (10.89)		26.37 (9.71)		0.13	[−0.52, 0.79]	
Pre-treatment	22.06 (10.78)	−0.51	22.32 (10.46)	−0.39	−0.02	[−0.68, 0.63]	
Post-treatment	15.65 (14.18)[a]	−0.94	10.11 (7.71)[a]	−1.82	0.48	[−0.18, 1.15]	
6-month follow-up	13.59 (11.89)[a]	−1.21	9.42 (8.51)[a]	−1.82	0.40	[−0.26, 1.06]	
12-month follow-up	13.95 (11.81)[a]	−1.19	10.59 (9.82)[a]	−1.58	0.30	[−0.35, 0.96]	
State Trait Anxiety Inventory – Trait (STAI-T)							
Baseline	50.06 (10.76)		54.47 (9.00)		−0.44	[−1.10, 0.22]	
Pre-treatment	50.18 (13.11)	0.01	51.32 (8.61)	−0.35	−0.10	[−0.76, 0.55]	
Post-treatment	43.71 (13.93)	−0.50	42.16 (10.19)[a]	−1.25	0.13	[−0.53, 0.78]	
6-month follow-up	42.65 (13.40)[a]	−0.60	41.93 (10.92)[a]	−1.23	0.06	[−0.60, 0.71]	
12-month follow-up	43.18 (14.14)	−0.53	41.21 (10.65)[a]	−1.32	0.15	[−0.50, 0.81]	
State Trait Anxiety Inventory – State (STAI-S)							
Baseline	44.41 (9.62)		43.11 (8.66)		0.14	[−0.52, 0.80]	
Pre-treatment	41.35 (11.29)	−0.28	42.16 (8.76)	−0.11	−0.08	[−0.73, 0.58]	
Post-treatment	36.65 (11.96)	−0.70	35.42 (10.15)[a]	−0.80	0.11	[−0.55, 0.76]	
6-month follow-up	35.82 (9.75)[a]	−0.87	33.53 (9.74)[a]	−1.02	0.23	[−0.43, 0.89]	
12-month follow-up	36.59 (14.31)	−0.63	38.37 (10.95)	−0.47	−0.14	[−0.79, 0.52]	
Beck Depression Inventory II (BDI-II)							
Baseline	18.24 (10.28)		21.89 (7.78)		−0.40	[−1.06, 0.27]	
Pre-treatment	14.59 (9.27)	−0.36	19.00 (8.99)	−0.34	−0.47	[−1.14, 0.19]	
Post-treatment	9.71 (9.25)[a]	−0.85	10.58 (9.29)[a]	−1.29	−0.09	[−0.75, 0.56]	
6-month follow-up	9.71 (8.77)[a]	−0.87	11.32 (8.06)[a]	−1.31	−0.19	[−0.84, 0.47]	
12-month follow-up	10.18 (11.13)[a]	−0.73	11.16 (8.99)[a]	−1.25	−0.10	[−0.75, 0.56]	
Quality of Life Inventory (QoLI)[b]							
Baseline	1.34 (1.52)		1.08 (1.26)		−0.19	[−0.47, 0.84]	
Pre-treatment	1.19 (1.18)	0.11	0.90 (1.52)	0.13	−0.21	[−0.44, 0.87]	
Post-treatment	1.64 (1.33)	−0.20	1.56 (1.65)	−0.33	−0.05	[−0.61, 0.70]	
6-month follow-up	1.87 (1.52)	−0.34	1.55 (1.62)	−0.32	−0.20	[−0.46, 0.85]	
12-month follow-up	1.85 (1.66)	−0.31	1.54 (1.84)	−0.29	−0.17	[−0.48, 0.83]	

Notes: MANOVAs assess combined effects of the different groups of outcome measures. Significant results are further assessed with subsequent ANOVAs for the implicated outcome measures. Effect sizes for each treatment are reported with reference to baseline values. All effect sizes are calculated with pooled standard deviations. Between-groups effects are reported with positive values indicating effect in favour of CBT. For the ANOVAs, degrees of freedom were corrected using Greenhouse-Geisser estimates. * = $p < .05$. [a]Significant within-group effect from Baseline to corresponding time-point (Bonferroni-corrected). [b]Valens of effect sizes were changed to be in concordance with the rest of the Table.

measures MANOVAs were performed for the groups on the primary outcome measures, using Pillai's trace (*V*) to evaluate the homogeneity of variance-covariance matrices. The combined analysis of the four inventories (MI-Alone, MI-Accompanied, ACQ and BSQ) revealed a significant effect of time from baseline to the 12-month follow-up, $F(16,544) = 7.28$, $p < .01$. There was also a significant Time×Group interaction ($p = .042$), which indicates that the two treatment groups differed over time on the primary outcome measures. Subsequently conducted repeated measures ANOVAs indicated that there was a significant main effect of
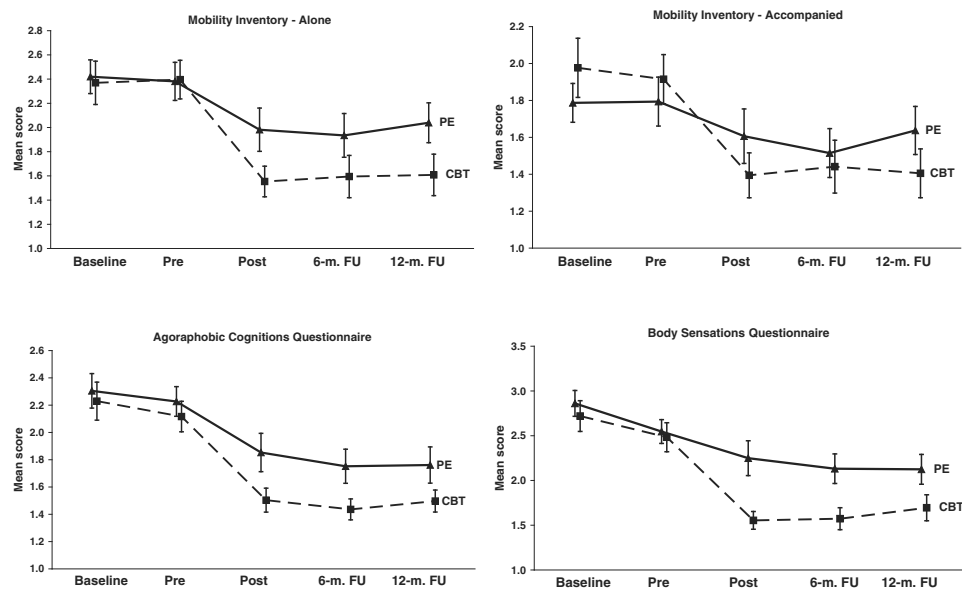
Figure 2. Physical exercise (PE) and CBT compared over time on the primary outcome inventories. Bars represent standard error of the mean

time for all four inventories: MI-Alone, $F(2.28, 77.61) = 36.44$, $p < .01$; MI-accompanied, $F(2.19, 74.28)$, $p < .01$; ACQ, $F(2.11, 71.62) = 31.88$, $p < .01$; and BSQ, $F(2.81, 95.43) = 41.69$, $p < .01$. For the BSQ there was also a significant effect of the intervention regardless of time, $F = 4.46$, $p = .042$. There was a significant Time×Group effect for the MI-Alone ($p = .023$), MI-Accompanied ($p = .021$) and BSQ ($p = .010$), but not for the ACQ ($p = .323$). See Figure 2 for a graphical presentation of these results.

A doubly two-way repeated measures MANOVA was also performed on the single-item self-report scales of panic frequency and panic-related distress and disability. This revealed a significant effect of time, $F(8, 272) = 15.48$, $p < .01$, whereas no significant Time×Group effect was found in the combined analysis of these two measures.

An equivalent assessment was performed on the single-item clinician-rated measures of panic frequency and panic-related distress and disability. A significant effect of time was found, $F(6, 204) = 16.69$, $p < .01$. The Time×Group interaction was however not significant.

Secondary outcome measures. The corresponding MANOVA performed on the five secondary outcome measures revealed a significant effect of time, $F(20,540) = 5.21$, $p < .01$. However, the Time×Group interaction was not significant, and consequently no further analysis was performed on the separate inventories. It should be noted that the five scales correlated highly for each assessment point, with the mean intercorrelations between the measures of .57, .57, .71, .70, and .72 for baseline, pre-treatment, post-treatment, 6-month follow-up and 12-month follow-up respectively.

Within-groups effects. Within-groups effects were assessed by post-hoc tests comparing baseline levels with levels at other points in time, with results reported in detail in Table 2. For the CBT group, these analyses yielded significant results ($p < .05$) at the 12-month follow-up on all measures except the Quality of Life scale and the STAI-S. For the physical exercise-group, the equivalent analyses yielded significant results at the 12-month follow-up on the following measures: the MI-Alone, the ACQ, the BSQ, both the self-rated- and the clinician-rated version of panic-related distress and disability, the BAI and the BDI-II.

Effect sizes. Both within- and between-groups effect sizes are summarized in detail in Table 2. Between-groups effect sizes at the 12-month follow-up were in favour of the CBT-intervention on all outcome measures, except for the STAI-S, the BDI-II and the Quality of Life Inventory. The within-groups effect sizes for physical exercise were found to be moderate on the MI-Alone on all assessments following treatment termination. On the MI-Accompanied the equivalent assessments showed small to moderate effects, whereas they were all large on the ACQ and the BSQ. For CBT all effects were large on these outcome measures for the corresponding assessments. Large effects were found on the reduction of panic-related distress and disability for both treatments on all assessments following the treatment period, although the effect sizes were considerably larger for CBT. On panic frequency moderate effects were found for physical exercise, whereas effects were large for CBT. Self-report and clinician-rated assessments were consistent for both panic-related distress and disability and panic frequency. Large effects were also found for both interventions on the BAI and the BDI-II for all assessments following treatment termination. The interventions differed on the STAI-T where effects were large for CBT and moderate for physical exercise. Both interventions yielded moderate to large effects on the STAI-S, and small effects on the QoLI.

Reliable and clinically significant change. Figure 3 displays the differences in both reliable and clinically significant change between the two interventions. While 63.2% of patients treated by CBT had a reliable improvement, this applied only to 35.3% of the patients in the physical exercise group. However, this difference was not significant ($p = .181$, two-sided Fisher's exact test). When the additional criterion of a clinically significant change was added, 52.6% of the CBT group fulfilled these criteria as compared to only 11.8% of the physical exercise group. This difference was significant ($p = .014$, two-sided Fisher's exact test). When adding the criterion of freedom from panic attacks, 11.8% of the physical exercise-group still fulfilled these criteria. For the CBT group, the proportion of participants satisfying the criteria was reduced to 47.4%. However, the difference between the two interventions remained significant ($p = .031$, two-sided Fisher's exact test).

Treatment-seeking and change in psychotropic medication at the 12-month follow-up Of the participants that completed the 12-month follow-up assessment, results showed that seven participants (50%) in the physical exercise-group had entered new psychotherapy directed at their panic, as compared to none of the participants in the CBT group ($p = .002$, two-sided Fisher's exact test). Furthermore, three participants in the physical exercise-group had either increased or been prescribed a new psychotropic medication directed at their panic, as compared to one in the CBT group ($p = .33$, two-sided Fisher's exact test).
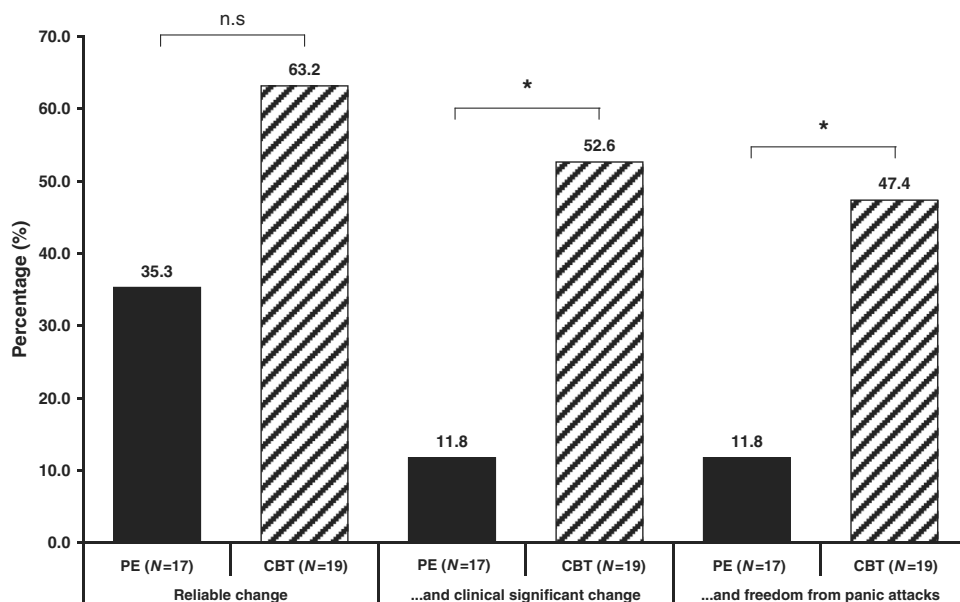
Figure 3. Reliable and clinically significant change from baseline to 12 month follow-up. Criteria are applied to MI – Alone, and are accumulated stepwise. Freedom from panic is based on the score on the panic frequency-item in the self-report measure Panic Attack Scale. PE = physical exercise. ∗ = $p <$ .05 (two-sided Fisher's exact test)

CBT adherence and competence

The quality of the CBT was assessed by CTACS and yielded a mean score of 4.7 for competence, 4.3 for adherence and 4.8 for the item that specifically assessed the overall performance of the therapist. Barber et al. (2003) reported mean values of 3.4 and 3.8 for adherence and competence, respectively, for cognitive therapy when assessing the criterion validity of the scale. As both the raters and the therapist in the study by Barber et al. were carefully selected and highly competent, the reported mean values in the current study suggest that both the content and delivery of the CBT were highly satisfactory.

Change in physical fitness following treatment

As shown in Table 1, there was no difference between the two interventions on the 6-minute walk test prior to treatment. Following the treatment-period, the physical exercise-group walked significantly longer than the CBT group, $t(34) = 2.15$, $p = .039$, on this test.

Missing data

Single items missing within inventories amounted to 0.12% of the total number of observations. Total or mean scores on primary and secondary outcome measures that were

missing due to lacking responses from participants or administrative error amounted to 7.14% of all observations.

Discussion

Both group CBT and group physical exercise were found to be effective for the treatment of PD, as assessed by primary and secondary outcome measures in this study. However, the effects were both greater and more consistent for CBT than for physical exercise. The most salient differences between the two treatments were seen on the primary outcome measures, and in particular on the inventories that measured agoraphobic avoidance (MI), fear of fear (ACQ) and fear of bodily sensations (BSQ). When the effects of the treatments were assessed in a combined analysis of these inventories, CBT was found to perform significantly better. With these inventories considered separately, CBT performed significantly better over time on all measures with the exception of the one assessing fear of fear. Physical exercise yielded large effects on both fear of fear and bodily sensations, as did CBT. However, the between-groups effect sizes indicated a larger effect in favour of CBT on bodily sensations, and especially at the post-treatment assessment. The effects on avoidance were smaller for both interventions. Still, they were large for CBT, while moderate and small for physical exercise on avoidance when alone and when accompanied, respectively. Both interventions reduced the patients' experienced distress or disability related to their PD. However, CBT appeared to do this more effectively. Furthermore, physical exercise did not significantly reduce the number of panic attacks, while CBT did. As the presence of recent panic attacks was not an inclusion criterion, not all participants experienced frequent attacks, thus the potential for changes in this domain was limited.

The analysis of the secondary outcome-measures indicated no difference between the two interventions over time. The results suggested that both interventions had the capacity to significantly reduce both symptoms of depression and general anxiety as measured by the BAI. The effects of treatment were generally lower as measured by the STAI-T and the STAI-S, as compared to what was found for the BAI. This is probably due to the fact that BAI, to a greater extent than the STAI-T and the STAI-S, reflects panic-related anxiety symptoms (Cox, Cohen, Direnfeld and Swinson, 1996).

Overall, we conclude that although physical exercise was associated with improvements on several outcome measures, CBT appears to yield a better treatment response. Improvement was also assessed in terms of reliable and clinically significant change (Jacobson and Truax, 1991). In the current study, CBT appeared to be more effective than physical exercise concerning clinically significant improvement. This is corroborated by the treatment-seeking behaviour displayed by participants within the 12 months following treatment termination.

Considering the results from the studies by Broocks et al. (1998) and Clark et al. (1994), one could argue that the results of the current study were to be expected. Broocks et al. found that exercise was outperformed by tricyclic antidepressants, whereas Clark et al. (1994) found cognitive therapy to be superior to a comparable antidepressant (Modigh, Westberg and Eriksson, 1992). These two studies are similar in methodology to the current study, and especially in terms of the treatment content. It is therefore also relevant to compare the results from the current study against what was found in these studies to assess the representativeness and validity of the results in the current study.

The current study compares well with the within-groups effect sizes that were found in the two above mentioned studies on physical exercise and cognitive therapy both in terms of primary (Bandelow, 1995) and secondary outcome measures. Effect sizes in the current study are slightly lower for both physical exercise and CBT compared to Broocks et al. (1998) and Clark et al. (1994), respectively. This could indicate that treatments in the current study were less effective. However, it could also reflect differences in sample characteristics. The sample of the current study had a longer mean duration of illness; 10.1 years compared to 3.1 (Broocks et al., 1998) and 3.3 (Clark et al., 1994). For the secondary outcome measures in the study by Broocks et al., only completer data were presented, thus possibly yielding a somewhat larger effect than that observed with an ITT-sample.

The current study also highlights an important challenge of studying the emotional effects of physical exercise, namely that of controlling the delivered dose (duration and intensity) of exercise. Previous studies with clinical samples (Broocks et al., 1998; Martinsen et al., 1989; Sexton et al., 1989) exerted relatively little control over the intensity during the exercise sessions. Subjects in these studies either walked or ran, alone or in groups. This contrasts with the more recent studies on anxiety sensitivity (Broman-Fulks et al., 2004; Broman-Fulks and Storey, 2008; Smits et al., 2008), where the delivered dose of exercise was carefully controlled through the use of treadmills. These studies do however also differ in terms of duration of the interventions, where subjects in the clinical studies performed the exercises for more than 8 weeks, whereas the subjects in the studies on anxiety sensitivity performed the exercises for 2 weeks. Previous research has indicated that the treatment should last at least 12 weeks to ensure reductions in trait-anxiety (Petruzzello et al., 1991), and Broocks et al. (1998) also suggested that a treatment-program longer than their 10 weeks may be necessary for the physical exercise to be fully effective. Drop-out has been a challenge in studies with clinical samples (Broocks et al., 1998; Meyer and Broocks, 2000; Sexton et al., 1989), and longer interventions naturally increase the likelihood for subjects to drop-out. Drop-out is a threat to the external validity of studies, as it questions the applicability of the intervention to the relevant population. Because the physical exercise in the current study took place three times per week over a period of 12 weeks, more varied treatment content was adopted to try to alleviate drop-out from the treatment. However, this allowed for less control over the delivered dose of exercise, especially with regard to the intensity of the exercise. As described in the Methods section, the delivered dose of physical exercise was controlled through the adherence to a detailed and structured treatment manual, and by having the exercise being performed under supervision on the site. In addition to this, intensity was monitored with a fitness heart rate monitor during the sessions intended to increase aerobic capacity. However, the importance of achieving increased aerobic capacity for participants or the necessity for participants to perform exercise at a high intensity during treatment sessions is still somewhat unclear. Though it has been shown that high-intensity exercise can reduce anxiety sensitivity better than low-intensity exercise (Broman-Fulks et al., 2004; Esquivel et al., 2008), other studies have indicated that increased aerobic fitness is not necessary to achieve reductions in anxiety and that resistance training or low-intensity exercise can provide anxiolytic effects (Martinsen et al., 1989; Meyer and Broocks, 2000; Sexton et al., 1989; Strohle et al., 2009). Regarding the current study, it is therefore not clear whether the subjects experienced reduced anxiety levels as a result of changes in aerobic capacity, or as a result of mere participation in 12 weeks of group physical exercise. Generally, group interventions are known to be associated with therapeutic benefits due to, for example, social support, normalization, peer

feedback and role modelling (Yalom and Leszcz, 2005). As such, it is possible that the use of group format in the present study could have influenced the treatment effects in both conditions by the above mentioned mechanisms.

The two interventions differed in frequency of delivery, as physical exercise was performed three times per week and CBT was performed once a week. As both treatments ran for 12 weeks, this means that the physical exercise-group received three times as many sessions. However, the two treatments were considered to represent two different packages of treatment in the current study. Adhering to the recommendations for the duration and frequency of physical exercise produced a format that could not readily be transferred to the standard methods of CBT treatment.

The current study had relatively fewer exclusion criteria than has been the case in comparable studies (Broocks et al., 1998; Clark et al., 1994). This was preferred with reference to the increased external validity this provides. Greater clinical realism in this field has been encouraged (Lawlor and Hopker, 2001; Salmon, 2001) and, compared to previous studies, the most noteworthy difference in the inclusion and exclusion criteria of the current study is the omission of the requirement of participants to have recently experienced panic attacks. As spontaneous and frequent panic attacks are known to decrease with time (American Psychiatric Association, 2000), this omission allows for a sample that has experienced chronic PD. This could obviously influence effect sizes on some of the measures as well as the rates of participants acquiring clinically significant change. However, the relatively large difference between the two interventions in terms of reliable and clinically significant change remains noteworthy.

The primary aim of this study was to evaluate the long-term clinical efficacy of physical exercise by comparing it with CBT, an established treatment for PD. Frequent calls have been made for the controlled assessment of the anxiolytic effects of physical exercise compared to an established treatment option within a relevant clinical context. Results here suggest that group CBT is more effective for treating PD than physical exercise performed in groups. However, physical exercise in itself yielded large and significant effects on the majority of both primary and secondary outcome measures at both post-treatment and the follow-up assessments. The current study provides insights into the long-term clinical effects of both physical exercise and CBT. Regarding physical exercise, a controlled assessment of these effects has not previously been provided. Both Figure 2 and Table 2 suggest that these effects do not diminish within the first 12 months. However, the assessment of the reliable and clinically significant change does suggest that clinical efficacy should be further explored before physical exercise can be recommended as a treatment for PD. This is emphasized by the differences between the two interventions on treatment-seeking behaviour following treatment termination.

Because physical exercise and CBT might improve panic symptoms by different mechanisms, future studies should further investigate whether physical exercise combined with CBT may yield greater improvements than CBT alone. Smits et al. (2008) found that the effect of physical exercise was not enhanced for subjects with increased anxiety sensitivity by the addition of a cognitive restructuring-component. Still, other efforts have been made to combine physical exercise with CBT as a treatment option for anxiety disorders, including PD. Two separate pilot studies have been conducted (Cromarty, Robinson, Callcott and Freeston, 2004; Merom et al., 2008), and both conclude that such efforts are acceptable to patients and that further evaluations should be carried out. Further research should also aim to reveal

whether physical exercise might work better than or as well as CBT for subgroups of patients suffering from PD.

Strengths and limitations

In the present study, several steps were taken to ensure the integrity of the interventions. Both treatment options were delivered by experienced therapists and according to a treatment manual. For the CBT, the treatment integrity was also evaluated by an external observer. An additional strength of the present study is that both treatments were highly attended and had low rates of attrition. All patients were assessed twice before the onset of treatment, and results indicated that no spontaneous recovery had taken place for any of the outcome measures of either group. Considering the mean duration of illness and the degree of co-morbidity, the current sample appears to represent the general population of PD patients. The treatments were analyzed in accordance with the principle of intention-to-treat, thus strengthening the external validity of the results. To our knowledge, the present study is the first RCT with long-term follow-up assessments that has investigated the clinical effects of physical exercise on PD.

In terms of study limitations, it should be noted that the therapist ratings were not blinded. The impact of this on the current study is considered small, as it only applies to two scales, which were also responded to by way of participant self-report. A more objective manipulation check than the 6-minute walk test (6MWT) should be preferred. However, considering the recent findings on the acute anti-panic effects of a single bout of physical exercise (Esquivel et al., 2008; Strohle et al., 2009), it is unclear what impact a more physically demanding test could have on treatment outcome. In terms of treatment content, the current study emphasized an ecological approach where two different treatment packages were compared. Although our power analysis indicated that we included a sufficient number of patients in the present study, certain non-significant effects might have been significant if a larger sample size had been used. Even though CBT is considered an evidence-based treatment for PD, the lack of a no treatment control group in the present study limits the ability to draw conclusions regarding the overall effectiveness of both interventions, as other explanations, such as statistical regression or demand effects, cannot be ruled out.

References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorder (4th ed., text rev.)*. Washington, DC: American Psychiatric Association.

Bandelow, B. (1995). Assessing the efficacy of treatments for panic disorder and agoraphobia. II. The Panic and Agoraphobia Scale. *International Clinical Psychopharmacology, 10*, 73–82.

Barber, J. P., Liese, B. S. and Abrams, M. J. (2003). Development of the Cognitive Therapy Adherence and Competence Scale. *Psychotherapy Research, 13*, 205–221.

Beck, A. T. and Steer, R. A. (1993). *Beck Anxiety Inventory Manual*. San Antonio, TX: The Psychological Corporation.

Beck, A. T., Steer, R. A. and Brown, G. K. (1996). *Beck Depression Inventory-II Manual*. San Antonio, TX: The Psychological Corporation.

Broman-Fulks, J. J., Berman, M. E., Rabian, B. A. and Webster, M. J. (2004). Effects of aerobic exercise on anxiety sensitivity. *Behaviour Research and Therapy, 42*, 125–136.

Broman-Fulks, J. J. and Storey, K. M. (2008). Evaluation of a brief aerobic exercise intervention for high anxiety sensitivity. *Anxiety Stress and Coping, 21*, 117–128.

Broocks, A., Bandelow, B., Pekrun, G., George, A., Meyer, T., Bartmann, U., et al. (1998). Comparison of aerobic exercise, clomipramine, and placebo in the treatment of panic disorder. *American Journal of Psychiatry, 155*, 603–609.

Chambless, D. L., Caputo, G. C., Bright, P. and Gallagher, R. (1984). Assessment of fear of fear in agoraphobics: the body sensations questionnaire and the agoraphobic cognitions questionnaire. *Journal of Consulting and Clinical Psychology, 52*, 1090–1097.

Chambless, D. L., Caputo, G. C., Jasin, S. E., Gracely, E. J. and Williams, C. (1985). The Mobility Inventory for Agoraphobia. *Behaviour Research and Therapy, 23*, 35–44.

Clark, D. M. (1986). A cognitive approach to panic. *Behaviour Research and Therapy, 24*, 461–470.

Clark, D. M., Salkovskis, P. M., Hackmann, A., Middleton, H., Anastasiades, P. and Gelder, M. (1994). A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. *The British Journal of Psychiatry, 164*, 759–769.

Clark, D. M., Salkovskis, P. M., Hackmann, A., Wells, A., Ludgate, J. and Gelder, M. (1999). Brief cognitive therapy for panic disorder: a randomized controlled trial. *Journal of Consulting and Clinical Psychology, 67*, 583–589.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cox, B. J., Cohen, E., Direnfeld, D. M. and Swinson, R. P. (1996). Does the Beck Anxiety Inventory measure anything beyond panic attack symptoms? *Behaviour Research and Therapy, 34*, 949–954.

Crapo, R. O., Casaburi, R., Coates, A. L., Enright, P. L., MacIntyre, N. R., McKay, R., et al. (2002). ATS statement: guidelines for the six-minute walk test. *American Journal of Respiratory and Critical Care Medicine, 166*, 111–117.

Cromarty, P., Robinson, G., Callcott, P. and Freeston, M. (2004). Cognitive therapy and exercise for panic and agoraphobia in primary care: pilot study and service development. *Behavioural and Cognitive Psychotherapy, 32*, 371–374.

Esquivel, G., Diaz-Galvis, J., Schruers, K., Berlanga, C., Lara-Munoz, C. and Griez, E. (2008). Acute exercise reduces the effects of a 35% CO_2 challenge in patients with panic disorder. *Journal of Affective Disorders, 107*, 217–220.

Esquivel, G., Schruers, K., Kuipers, H. and Griez, E. (2002). The effects of acute exercise and high lactate levels on 35% CO_2 challenge in healthy volunteers. *Acta Psychiatrica Scandinavica, 106*, 394–397.

Faul, F., Erdfelder, E., Lang, A. G. and Buchner, A. (2007). G∗Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

First, M. B., Spitzer, R. L., Gibbon, M. and Williams, J. B. W. (1995). *Structured Clinical Interview for DSM-IV Axis I Disorders-Patient Edition (SCID-I/P, Version 2.0)*. New York: Biometric Research Department, New York State Psychiatric Institute.

First, M. B., Spitzer, R. L., Gibbon, M., Williams, J. B. W. and Benjamin, L. (1994). *Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II, Version 2.0)*. New York: Biometrics Research Department, New York State Psychiatric Institute.

Frisch, M. B., Cornell, J., Villanueva, M. and Retzlaff, P. J. (1992). Clinical validation of the Quality of Life Inventory: a measure of life satisfaction for use in treatment planning and outcome assessment. *Psychological Assessment, 4*, 92–101.

Gould, R. A., Otto, M. W. and Pollack, M. H. (1995). A meta-analysis of treatment outcome for panic disorder. *Clinical Psychology Review, 15*, 819–844.

Haskell, W. L., Lee, I. M., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A., et al. (2007). Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Medicine and Science in Sports and Exercise, 39*, 1423–1434.

Herring, M. P., O'Connor, P. J. and Dishman, R. K. (2010). The effect of exercise training on anxiety symptoms among patients: a systematic review. *Archives of Internal Medicine, 170*, 321–331.

Jacobson, N. S. and Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19.

King, S., Wessel, J., Bhambhani, Y., Maikala, R., Sholter, D. and Maksymowych, W. (1999). Validity and reliability of the 6 minute walk in persons with fibromyalgia. *Journal of Rheumatology, 26*, 2233–2237.

Lawlor, D. A. and Hopker, S. W. (2001). The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. *British Medical Journal, 322*, 763–767.

Long, B. C. and van Stavel, R. (1995). Effects of exercise training on anxiety: a meta-analysis. *Journal of Applied Sport Psychology, 7*, 167–189.

Martinsen, E. W., Hoffart, A. and Solberg, O. Y. (1989). Aerobic and non-aerobic forms of exercise in the treatment of anxiety disorders. *Stress Medicine, 5*, 115–120.

Martinsen, E. W., Sandvik, L. and Kolbjørnsrud, O. B. (1989). Aerobic exercise in the treatment of nonpsychotic mental disorders: an exploratory study. *Nordic Journal of Psychiatry, 43*, 521–529.

McHugh, R. K., Otto, M. W., Barlow, D. H., Gorman, J. M., Shear, M. K. and Woods, S. W. (2007). Cost-efficacy of individual and combined treatments for panic disorder. *Journal of Clinical Psychiatry, 68*, 1038–1044.

Merom, D., Phongsavan, P., Wagner, R., Chey, T., Marnane, C., Steel, Z., et al. (2008). Promoting walking as an adjunct intervention to group cognitive behavioral therapy for anxiety disorders: a pilot group randomized trial. *Journal of Anxiety Disorders, 22*, 959–968.

Meyer, T. and Broocks, A. (2000). Therapeutic impact of exercise on psychiatric diseases: guidelines for exercise testing and prescription. *Sports Medicine, 30*, 269–279.

Modigh, K., Westberg, P. and Eriksson, E. (1992). Superiority of clomipramine over imipramine in the treatment of panic disorder: a placebo-controlled trial. *Journal of Clinical Psychopharmacology, 12*, 251–261.

National Institute for Health and Clinical Excellence (2004). *Anxiety: management of anxiety (panic disorder, with or without agoraphobia, and generalised anxiety disorder) in adults in primary, secondary and community care*. London: NICE.

Nes, B. M., Janszky, I., Wisløff, U., Støylen, A. and Karlsen, T. (2012). Age-predicted maximal heart rate in healthy subjects: the HUNT Fitness Study. *Scandinavian Journal of Medicine and Science in Sports*. Advance online publication. doi:10.1111/j.1600-0838.2012.01445.x

Orwin, A. (1974). Treatment of a situational phobia: a case for running. *British Journal of Psychiatry, 125*, 95–98.

Petruzzello, S. J., Landers, D. M., Hatfield, B. D., Kubitz, K. A. and Salazar, W. (1991). A meta-analysis on the anxiety-reducing effects of acute and chronic exercise: outcomes and mechanisms. *Sports Medicine, 11*, 143–182.

Pollock, M. L., Gaesser, G. A., Butcher, J. D., Despres, J. P., Dishman, R. K., Franklin, B. A., et al. (1998). ACSM position stand: the recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. *Medicine and Science in Sports and Exercise, 30*, 975–991.

Salmon, P. (2001). Effects of physical exercise on anxiety, depression, and sensitivity to stress: a unifying theory. *Clinical Psychology Review, 21*, 33–61.

Schmidt, N. B., Lerew, D. R. and Jackson, R. J. (1999). Prospective evaluation of anxiety sensitivity in the pathogenesis of panic: replication and extension. *Journal of Abnormal Psychology, 108*, 532–537.

Sexton, H., Maere, A. and Dahl, N. H. (1989). Exercise intensity and reduction in neurotic symptoms. *Acta Psychiatrica Scandinavica, 80*, 231–235.

Smits, J. A. J., Berry, A. C., Rosenfield, D., Powers, M. B., Behar, E. and Otto, M. W. (2008). Reducing anxiety sensitivity with exercise. *Depression and Anxiety, 25*, 689–699.

Smits, J. A. J., Meuret, A. E., Zvolensky, M. J., Rosenfield, D. and Seidel, A. (2009). The effects of acute exercise on CO_2 challenge reactivity. *Journal of Psychiatric Research, 43*, 446–454.

Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., Vagg, P. R. and Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory (Form Y)*. Palo Alto, CA: Consulting Psychologists Press.

Stein, J. M., Papp, L. A., Klein, D. F., Cohen, S., Simon, J., Ross, D., et al. (1992). Exercise tolerance in panic disorder patients. *Biological Psychiatry, 32*, 281–287.

Strohle, A., Feller, C., Onken, M., Godemann, F., Heinz, A. and Dimeo, F. (2005). The acute antipanic activity of aerobic exercise. *American Journal of Psychiatry, 162*, 2376–2378.

Strohle, A., Graetz, B., Scheel, M., Wittmann, A., Feller, C., Heinz, A., et al. (2009). The acute antipanic and anxiolytic activity of aerobic exercise in patients with panic disorder and healthy control subjects. *Journal of Psychiatric Research, 43*, 1013–1017.

Telch, M. J., Lucas, J. A., Schmidt, N. B., Hanna, H. H., Jaimez, T. L. and Lucas, R. A. (1993). Group cognitive-behavioral treatment of panic disorder. *Behaviour Research and Therapy, 31*, 279–287.

Wedekind, D., Broocks, A., Weiss, N., Engel, K., Neubert, K. and Bandelow, B. (2010). A randomized, controlled trial of aerobic exercise in combination with paroxetine in the treatment of panic disorder. *World Journal of Biological Psychiatry, 11*, 904–913.

Wipfli, B. M., Rethorst, C. D. and Landers, D. M. (2008). The anxiolytic effects of exercise: a meta-analysis of randomized trials and dose-response analysis. *Journal of Sport and Exercise Psychology, 30*, 392–410.

Yalom, I. D., and Leszcz, M. (2005). *Theory and Practice of Group Psychotherapy (5th ed.)*. New York: Basic Books.